PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00, 1/20 — A1

(11) International Publication Number: WO 88/ 07079

(43) International Publication Date: 22 September 1988 (22.09.88)

(21) International Application Number: PCT/US88/00374

(22) International Filing Date: 8 February 1988 (08.02.88)

(31) Priority Application Number: 023,637

(32) Priority Date: 9 March 1987 (09.03.87)

(33) Priority Country: US

(71) Applicant: CETUS CORPORATION [US/US]; 1400 Fifty-Third Street, Emeryville, CA 94608 (US).

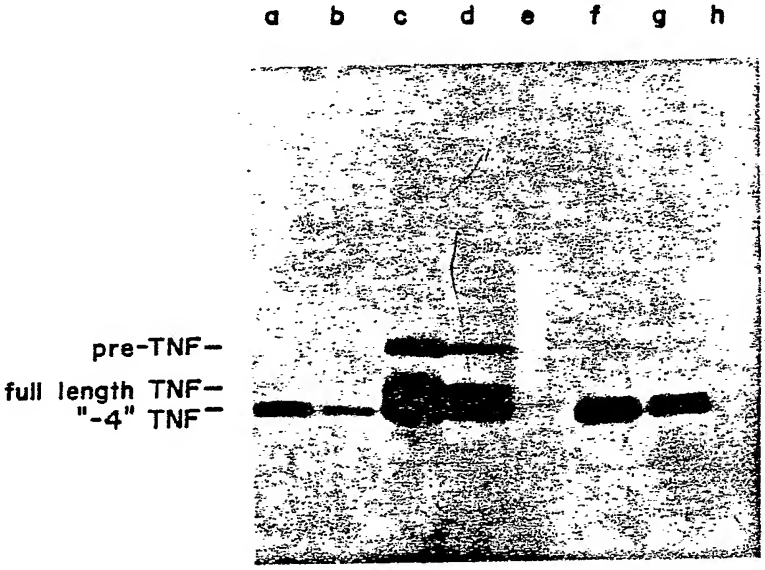
(72) Inventors: CHANG, Shing ; CHANG, Shen-Yung ; 2289 Manzanita Drice, Oakland, CA 94611 (US).

(74) Agent: HALLUIN, Albert, P.; Cetus Corporation, 1400 Fifty-Third Street, Emeryville, CA 94608 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.

(54) Title: EXPRESSION OF HETEROLOGOUS GENES IN STREPTOMYCES SPECIES



(57) Abstract

Vectors effective for the expression and secretion of heterologous genes in streptomycetes are disclosed. Such vectors comprises a plasmid replicable in *Streptomyces*, which comprises a *Streptomyces* promoter and a DNA sequence encoding a *Streptomyces* signal sequence operably linked to, and under control of the promoter, which DNA sequence encoding a *Streptomyces* signal sequence may be operably linked to a heterologous gene encoding a desired protein so that the *Streptomyces* signal sequence and heterologous protein are expressed, and the desired protein is secreted under control of the *Streptomyces* signal sequence. Various *Streptomyces* promoters and secretion signal sequences may be used in the invention, and DNA sequences comprising the promoters operably linked to DNA sequences encoding *Streptomyces* signal sequences, are disclosed as part of the invention. Vectors comprising the *Streptomyces aph, ermE, ermEa1* and modifications thereof, controlling the expression of the *amy* and ORF438 signal sequences, which lead to secretion of desired heterologous proteins when transformed into *Streptomyces* hosts, are exemplified.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

Expression of heterologous genes in streptomyces species.

This invention comprises novel recombinant DNA cloning vectors suitable for the expression and secretion of heterologous proteins in Streptomyces, Streptomyces hosts transformed with such vectors, and methods for the secreted production of heterologous proteins in Streptomyces.

Most of the streptomycetes bacteria are soil microorganisms. They synthesize secreted proteins and release them into surrounding environments. Up to the present, the genus Bacillus has attracted most of the interest as a Gram-positive host system for the secretion of heterologous proteins using recombinant DNA methods. However, even when grown to high density, transformed Bacilli rarely produce no greater than 10 µg/ml of secreted heterologous mammalian proteins. Streptomyces, many strains of which are known in the pharmaceutical arts as the source of antibiotics, are an attractive host system for secreted expression of protein because they are known to secrete antibiotics.

In general, a number of cloning vectors for use in Streptomyces are known. In particular, U.S. Patent Nos. 4,513,085 and 4,513,086 disclose selectable recombinant DNA cloning vectors for use in Streptomyces. Generally, the foregoing patents describe recombinant DNA cloning vectors which comprise an origin of replication on a restriction fragment of the plasmid, together with one or more DNA segments that confer resistance to antibiotics. The resistance to antibiotics may be used as a selection marker by which transformants of the plasmids may be selected.

Neither of the patents however disclose the isolation of functional promoters or secretion signal sequences that may be used either alone or together to express or secrete particular heterologous proteins on a recombinant plasmid suitable for use in Streptomyces. European Patent Application (EPA) 84303755.7 published July 17, 1985

describes recombinant plasmids suitable for expressing foreign DNA sequences in Streptomyces in which the plasmid is characterized by a promoter of the Streptomyces aminoglycoside phosphotransferases. The expression of the eucaryotic protein bovine growth hormone as a non-secreted product in Streptomyces is described. The patent application asserts that the methods of the invention enable production and excretion of products in Streptomyces through recombinant DNA technology. In addition, there is a general description of an embodiment of an expression plasmid in which DNA encoding a signal sequence of a polypeptide or protein normally secreted from a Streptomyces host employed in the expression system is operatively linked to the Streptomyces aminoglycoside phosphotransferase (aph) promoter which is used as the expression control system, to achieve the predicted production of a putative fusion protein. The putative fusion protein is predicted to effect the transport of the fusion protein through the Streptomyces cell wall and the maturation of the desired protein by cleavage of the signal sequence. The disclosure however lacks the description of any suitable Streptomyces or other secretion signal and, furthermore, fails to show a single example in which such secreted expression is accomplished.

EPA 84303755.7 however does describe Streptomyces lividans 66 containing plasmid pIJ703 which was deposited on June 6, 1983 into the American Type Culture Collection under Accession No. 39378. pIJ703 is identical to pIJ702, which is described further hereinbelow, except that the fragments containing the gene coding for tyrosinase determinant (mel) and the gene coded for thiostreptone resistance (tsr) are transposed as described in Katz et al., J. Gen. Microbiol. 129:2703-14 (1983).

Gray et al., Gene 32:22-30 (1984) discloses the production of bovine growth hormone using a DNA sequence that, in the authors' estimation, probably contains the control regions of the S. fradiae aph gene which leads to the intracellular accumulation of a product that is shown to be a bovine growth hormone by immunoblotting and radioimmunoassay. The aph promoter, in plasmid pEC23, appears to be the same as that found in pEC24 of the previous reference.

Keiser et al. in Mol. Gen. Genet., 185:223-238 (1982), disclose a number of broad host range multicopy plasmids suitable for use as cloning vectors in Streptomyces. In particular, plasmid pIJ101 is disclosed. Katz et al. disclosed a number of plasmids derived from pIJ101 and in particular disclosed the construction of pIJ702 which has two markers, tsr and mel. pIJ702 is the same plasmid as pIJ703, ATCC #39378 except that the BclI fragment containing the tsr and mel genes are oriented in the opposite direction.

A number of Streptomyces signal peptide sequences are reported in the literature. The signal sequence peptides listed in Table I appear to be functional for the secretion of the homologous protein to which they are attached prior to secretion in nature. In general, these Streptomyces signal peptide sequences have a typical structure for signal peptides. This general structure is characterized by a short stretch of N-terminal residues containing several positively charged residues, followed by a segment of neutral mostly hydrophobic residues followed, frequently, by a proline residue at the -4 to -7 position. The signal peptidase usually cleaves after the Ala-x-Ala sequence at the -3 to -1 location, wherein x at -2 is an amino acid that may vary. In the signal sequences identified in Table I, for example, x may be proline, serine, histidine, arginine, or alanine. The +1 position is taken as the amino terminal residue of the mature protein.

TABLE I

Organism/Gene/Strain	Sequence
Streptomyces plicatus endo H	MFTPVRRRVRTAALALSAAAALVLGSTAASGASATPSPAPA/P/APAPVKQGPT
S. avidinii streptavidin	(MSH)MRKIVVAAIAVSLTTVSITASASA/DPSKDSKAQV
S. coelicolor: agarase	MVNRRDLIKWSAVALGAGAGLAGPAPAAHA/ADLEWEQYPV
S. lividans: β-galactosidase	MPHSPVSPAESPAPQPGRRRPVVSRRRLLE GGAAVLGALALSASPLTAQAAVRRAA/ADEPPEWNDF
S. antibioticus: ORF438	MPELTRRRALGAAAVVAAGVPLVALPAARA/DDRGHHTPEV
S. limosus: α-amylase	MARRLATASLAVLAAAATALTAPTPAAA/APPGAKDVTA

Streptomyces offer a number of advantages for the production of recombinant proteins and it would therefore be desirable to construct vectors suitable for secreting heterologous gene products in Streptomyces. Among the desirable characteristics of Streptomyces are the following, they are in general non-toxic and do not cause disease in man or animals. Furthermore, Streptomyces have been used in industrial fermentation processes to produce antibiotics for many years and as a result there is considerable knowledge and experience in large scale Streptomyces fermentations. Indeed, the economics of production of proteins with Streptomyces is so favorable that enzymes useful in food production may be produced as commodities using Streptomyces industrial fermentations and purifications. Streptomyces may be grown to high culture densities and will actively synthesize and secrete homologous products for several weeks after the completion of agitated growth. Despite the clear advantages of the use of Streptomyces for the production of recombinantly produced proteins, there has not, heretofore been reported an expression vector capable of secreting heterologous proteins of eukaryotic, particularly mammalian, viral or prokaryotic origin in Streptomyces.

In one aspect, the invention concerns recombinant vectors suitable for producing and secreting heterologous proteins in Streptomyces. In general, these recombinant secretion vectors comprise a plasmid replicable in Streptomyces which further comprises a promoter functional in Streptomyces and a DNA sequence encoding a secretion signal sequence that is operable in Streptomyces and may be operably linked to a heterologous gene under the control of said promoter. In a preferred embodiment, the DNA sequence encoding the secretion signal sequence further includes a site for insertion of a heterologous gene to produce a protein that is processed by the secretion signal. In general, high level promoters, preferably from Streptomyces organisms, that control the initiation of many mRNA transcripts of the gene which they control are preferred. Among the promoters that may be used in the secretion vectors according to the invention are the promoter of the aminoglycoside phosphotransferase gene (aph) promoter, the ermE promoter, and modified versions of the

ermE promoter which are described in further detail hereinbelow. The secretion signal sequences that may be used in the secretion vector, according to the invention, are engineered from proteins that are secreted by Streptomyces and include the secretion signal sequence of the α-amylase gene of Streptomyces limosus (amy signal sequence) and the signal sequence of the ORF438 gene (ORF438 signal sequence). Other exemplary Streptomyces signal sequences that may be useful in the vectors according to the invention include those found in Table I.

In another aspect the invention relates to a recombinant DNA sequence comprising a high level promoter operable in Streptomyces which is operably linked to a secretion signal sequence operable in Streptomyces. In a preferred embodiment, the secretion signal sequence will have at least one site for the insertion of a DNA sequence encoding a desired heterologous protein. The insertion site is postioned such that the signal sequence and heterologous gene product, when translated in a Streptomyces host, produce a protein that is processed, i.e., cleaved from the secretion signal sequence and preferably secreted by the Streptomyces host.

In another aspect of the invention, the invention relates to the foregoing secretion vectors into which a gene encoding a DNA sequence foreign to Streptomyces has been ligated so that a protein comprising the signal secretion signal sequence fused to the heterologous protein is produced and secreted.

In yet another aspect, the invention concerns Streptomyces transformed with the foregoing secretion vectors, which transformed Streptomyces secrete the product encoded by the foreign DNA sequence.

The invention will be better understood in connection with the following figures.

Figure 1 is the sequence of the ermE promoter as determined by Bibb et al., Gene, 38:215-226 (1985); Gene, 41:E357-368 (1986). The "TGG" sequence removed by site-specific mutagenesis to produce ermE is overlined with the triangle.

Figure 2 is the sequence of the aph promoter as determined by Bibb et al., Molec. Gen. Genet. 199:26-36 (1985).

Figure 3A is a composite restriction map of the fragments containing the amy gene which includes 2.3-kb SstI-EcoRI fragment from pIJ702.No5SL, and the 2.25-kb KpnI-KpnI fragment from pIJ2921.SL.

Figure 3B shows the position relative to Figure 3A of the transcription start site (wavy line) the direction of transcription (arrow), the coding sequences for the signal peptide (hatched box) and the mature amylase (open box).

Figure 3C shows the amy gene fragment (thin line) in plasmid pSYC1388. Sequence derived from pIJ702 is shown in wavy line.

Figure 4 shows the nucleotide sequence of the entire amy gene from S. limosus. The upstream open reading frame (ORF) protein is translated. The inverted repeat sequence is marked above by >>><<<. The transcription start site is identified by ! and the "-10" and "-35" regions in the promoter are identified by overlining. The deduced amino acid sequence of the putative preamylase protein is numbered from the initiation methionine. The arrow marks the junction of the amylase signal peptide and mature protein sequences.

Figure 5 is a schematic representation of the procedure used for producing the aph-amy signal sequence and ermE-amy signal sequence Streptomyces secretion vectors. In this figure, as in all of the subsequent figures, a restriction site followed by (r) means that the site was repaired by filling in the cut end using Klenow fragment and dNTPs. If followed by (S1), the cut end was trimmed with S1 exonuclease. If followed by a (t) 3' single-stranded cut ends were removed by the exonucleolytic action of the Klenow fragment. All three procedures yield blunt ends.

Figure 6 shows the primary nucleotide sequence of the S. antibioticus tyrosinase fragment. Upward arrow ( ) denotes potential signal sequence cleavage site in ORF438 gene product. Underlined nucleotide (nt) sequence denotes: nt 138-148 = -35 region; nt 163-169 = -10 region; nt 258-263 and nt 730-733 = RBS. Asterisks (***)

specify terminator codons of ORF438 and of mel genes. ORF438 extends from nt 270 to nt 707.

Figure 7 is a schematic representation of the procedure used for producing the aph-ORF secretion signal Streptomyces secretion vectors pSYC1289 and pSYC1290.

Figure 8 is a schematic representation of the procedure used for producing secreted tumor necrosis factor using aph-amy secretion signal and ermE-amy secretion signal Streptomyces secretion vectors.

Figure 9 shows a 12.5% SDS-PAGE of TCA precipitated culture supernatant from cultured Streptomyces lividans TK24 transformed with pSYC1414. 1.1 ml supernatant of cells grown in R6 medium were precipitated with final concentration of 20% TCA. Pellets were washed with acetone: ethanol (1:1) twice and resuspended in sample cracking buffer. Lane 1, ermE promoter- amy signal peptide - TNF (pSYC1414). Lane 2, negative control (pIJ702 with inactivated mel gene). Lane 3, low molecular weight marker.

Figure 10 is a schematic representation of the procedure used for producing E. coli alkaline phsophatase in Streptomyces using the aph-ORF secretion vector.

Figure 11 is a 12.5% SDS-PAGE of concentrated culture supernatant of pSYC1328 which comprised the aph promoter - ORF438 signal peptide - alkaline phosphatase fusion. Lane 1, low molecular weight marker. Lanes 2, 3 and 4, pM30 membrane concentrated from 0.3 ml, 0.6 ml and 1.5 ml supernatant, respectively. Lane 5, 20% TCA precipitated from 0.8 ml supernatant. Lane 6, 0.5 µg purified bacterial alkaline phosphatase (SIGMA).

Figure 12 shows the synthetic ltx gene. The single-stranded synthetic fragments are identified with double lines, and are numbered at the 5' ends. The expected amino acid sequence is shown in single-letter observations recommended by the IUPAC-IUB commission on Biochemical Nomenclature.

Figure 13 is a SDS-PAGE analysis of the extracellular proteins from TK24 (pSYC1483). Lanes (1), molecular weight standards,

as marked, in Kilodaltons (Kd); (2) and (3), extracellular proteins precipitated from cultures of two independent isolates on day 5; (4) and (5), samples from day 7; (6) sample from TK24. The arrow marks the 16-kd protein band.

Figure 14 is a schematic representation of the procedure used for producing the ermE-ORF Streptomyces secretion vectors used to express and secrete lymphotoxin.

Figure 15 is a schematic representation of one process used for producing the ermE1a-ORF Streptomyces secretion vector used to express and secrete TNF.

Figure 16 is an SDS-PAGE gel showing on analysis of secreted and cellular proteins produced by Streptomyces secretion vectors secreting TNF.

Figure 17 is a Western Blot of cellular and secreted proteins produced by Streptomyes secretion vectors secreting TNF.

A. Definitions

In the detailed description of the invention, the following terms are employed.

"Plasmid" as used herein means a covalently closed, generally double-stranded circular DNA sequence which comprises an intact replicon or replicating unit which is replicated in a host cell.

"Promoter" as used herein is meant one of a number of DNA sequences involved in the control of the initiation of transcription of a gene. In general, a promoter is a DNA sequence which is involved in the initiation of transcription of a DNA sequence encoding a protein. Promoters frequently have a number of distinct structural features and may, in addition, be regulated by operators which control the initiation of transcription or attenuators, which are involved in the sustaining of transcription. Most promoters include a Shine/Dalgarno sequence, however this sequence is not found in all promoters, as is known with respect to, for example, the aph promoter.

"Control sequences" as used herein refers to DNA sequences necessary for the expression of an operably linked coding sequence in the Streptomyces host organism. The control sequences which are suitable for Streptomyces include a promoter, a ribosome binding site and, in general, the Shine/Dalgarno sequence, although it is known, for example, that the aph promoter may operate and generally does operate in the absence of any definable Shine/Dalgarno sequence.

"Streptomyces host" as used herein is meant any strain of Streptomyces in which the recombinant vector according to the invention is functional. pIJ702 and derivatives thereof have been found to be functional in every strain of Streptomycetes into which they have been transformed. Among the Streptomyces hosts contemplated within the definition are the following: S. lividans, S. coelicolor A3(2), S. griseus, S. parvulus, S. albus G, S. vinaceus, S. acrimycinis, S. calvuligerus, S. limosus, S. rubiginosis, S. azureus, S. glaucenscens, S. rimosus, S. violaceotuber, S. kanamyceticus or any streptomycetes host in which the replicon of PIJ702 is functional. By "transformed" or "transformants" is meant host cells into which a recombinant vector has been introduced and which maintains the vector therein, usually in the presence of an agent, such as an antibiotic, the resistance to which is conferred by the transforming plasmid.

"Operably linked" as used herein refers to the juxtaposition of control sequences, such as a promoter ribosome binding site in a manner that the normal function of the components is maintained. Thus, a coding sequence "operably linked" to an encoded secretion signal sequence refers to a configuration wherein the coding sequence is joined to the signal sequence in such a manner that the signal peptide is processed by the host cell and the process protein secreted. A secretion signal sequence operably linked to a promoter is joined to the promoter in such a manner that the transcription and translation of the secretion signal sequence is controlled by the promoter, ribosome binding site, and Shine/Dalgarno sequence if required.

"Heterologous" when used herein in connection with DNA refers to DNA sequences that originate in an organism which is a different species and usually a different genus than the species or genus of the host cell into which the heterologous DNA is placed. With respect to DNA encoding a protein or heterologous gene, heterologous as used herein means DNA from an organism other than the genus Streptomyces. "Homologous" as used herein with respect to DNA encoding a protein means DNA from an organism of the genus Streptomyces.

B. General Methods for Carrying out the Invention

Transformations

Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described by Cohen, S. N., Proc. Natl. Acad. Sci. (USA) (1972) 69:2110, or the $RbCl_2$ method described in Maniatis, et al., Molecular Cloning: A Laboratory Manual (1982) Cold Spring Harbor Press, p. 254 was used for procaryotes or other cells which contain substantial cell wall barriers.

Growth of Streptomyces and molecular biological methods particularly useful in connection therewith are described in Hopwood et al., Genetic Manipulation of Streptomyces A Laboratory Manual (1985). The John Innes Foundation.

Vector Construction

Construction of suitable vectors containing the desired coding and control sequences employs standard ligation and restriction techniques which are well understood in the art. Isolated plasmids, DNA sequences, or synthesized oligonucleotides are cleaved, tailored, and religated in the form desired.

Site specific DNA cleavage is performed by treating with the suitable restriction enzyme (or enzymes) under conditions which are generally understood in the art, and the particulars of which are specified by the manufacturer of these commercialy available

restriction enzymes. See, e.g., New England Biolabs, Product Catalog. In general, about 1 µg of plasmid or DNA sequence is cleaved by one unit of enzyme in about 20 µl of buffer solutions. In the examples herein, typically, a 3-10 fold excess of restriction enzyme is used to insure complete digestion of the DNA substrate. Incubation times of about one hour to two hours at 37°C or other appropriate temperatures are workable, although variations can be tolerated. After each incubation, protein is removed by extraction with phenol/chloroform, and may be followed by ether extraction, and the nucleic acid recovered from aqueous fractions by precipitation with ethanol followed by running over a Sephadex G-50 spin column. If desired, size separation of the cleaved fragments may be performed by polyacylamide gel or agarose gel electrophoresis using standard techniques. A general description of size separations is found in Methods in Enzymology (1980) 65:499-560 or Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).

Restriction cleaved fragments may be blunt ended by treating with the large fragment of E. coli DNA polymerase I (Klenow) in the presence of the four deoxynucleotide triphosphates (dNTPs) using incubation times of about 15 to 25 minutes at 20 to 25°C in 50 mM Tris pH 7.6, 50 mM NaCl, 6 mM $MgCl_2$, 6 mM DTT and 5-10 µM dNTPs. The Klenow fragment fills in at 5' overhangs but chews back protruding 3' single strands, even though the four dNTPs are present. If desired, selective repair can be performed by supplying only one of the, or selected, dXTPs within the limitations dictated by the nature of the sticky ends. After treatment with Klenow, the mixture is extracted with phenol/chloroform and ethanol precipitated. Treatment under appropriate conditions with S1 nuclease results in hydrolysis of any single-stranded portion.

Synthetic oligonucleotides are prepared by the triester method of Matteucci, et al. (J. Am. Chem. Soc. (1981) 103:3185) or using commercially available automated oligonucleotide synthesizers. Kinasing of single strands for labeling is achieved using an excess, e.g., approximately 10 units of polynucleotide kinase to 10 pmole

substrate in the presence of 50 mM Tris, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol, 40 pmoles γ-32P-ATP (3000 Ci/mmole), 0.1 mM spermidine, 0.1 mM EDTA.

Ligations are performed in 15-30 µl volumes under the following standard conditions and temperatures: 20 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 33 µg/ml BSA, 10 mM-50 mM NaCl, and either 40 µM ATP, 0.01-0.02 (Weiss) units T4 DNA ligase at 0°C (for "sticky end" ligation) or 1 mM ATP, 0.3-0.6 (Weiss) units T4 DNA ligase at 14°C (for "blunt end" ligation). Intermolecular "sticky end" ligations are usually performed at 33-100 µg/ml total DNA concentrations (5-100 nM total end concentration). Intermolecular "blunt end" ligations (usually employing a 10-30 fold molar excess of linkers) are performed at 1 µM total ends concentration.

In vector construction employing "vector fragments", the vector fragment is commonly treated with bacterial alkaline phosphatase (BAP) in order to remove the 5' phosphate and prevent religation of the vector. BAP digestions are conducted at pH 8 in approximately 150 mM Tris, in the presence of Na^+ and Mg^{+2} using about 1 unit of BAP per µg of vector at 60°C for about one hour. Vector fragments subjected to this treatment are referred to herein as "bapped". If unkinased oligodeoxyribonucleotides are used however, the vector fragments are not "bapped". In order to recover the nucleic acid fragments, the preparation is extracted with phenol/chloroform and ethanol precipitated and desalted by application to a Sephadex G-50 spin column. Alternatively, religation can be prevented in vectors which have been double digested by additional restriction enzyme digestion of the unwanted fragments.

For portions of DNA which require sequence modifications, site specific primer directed mutagenesis is used. This is conducted using a synthetic oligonucleotide primer complementary to a single stranded phage DNA to be mutagenized except for limited mismatching, representing the desired mutation. Briefly, the synthetic oligonucleotide is used as a primer to direct synthesis of a strand complementary to the phage, and the resulting double-stranded DNA is

transformed into a phage-supporting host bacterium. Cultures of the transformed bacteria are plated in top agar, permitting plaque formation from single cells which harbor the phage.

Theoretically, 50% of the new plaques will contain the phage having, as a single strand, the mutated form; 50% will have the original sequence. The resulting plaques are hybridized with kinased synthetic primer at a temperature which permits hybridizaton of an exact match, but at which the mismatches with the original strand are sufficient to prevent hybridization. Plaques which hybridize with the probe are then picked, cultured, and the DNA recovered. Details of site specific mutation procedures are described below in specific examples.

Verification of Construction

In the constructions set forth below, correct ligations for plasmid construction are confirmed by first transforming E. coli strain MM294 obtained from E. coli Genetic Stock Center, CGSC 6135, or other suitable host with the ligation mixture. Successful transformants are selected by ampicillin, tetracycline or other antibiotic resistance or using other markers depending on the mode of plasmid construction, as is understood in the art. Plasmids from the transformants are then prepared according to the method of Clewell, D. B., et al., Proc. Natl. Acad. Sci. (USA) (1969) 62:1159, optionally following chloramphenicol amplification (Clewell, D. B., J. Bacteriol. (1972) 110:667). The isolated DNA is analyzed by restriction enzyme mapping and/or sequenced by the dideoxy method of Sanger, F., et al., Proc. Natl. Acad. Sci. (USA) (1977) 74:5463 as further described by Messing, et al., Nucleic Acids Res. (1981) 9:309, or by the method of Maxam, et al., Methods in Enzymology (1980) 65:499.

Hosts Exemplified

Host strains used in cloning and expression herein are as follows:

For cloning and sequencing, and for expression of construction under control of most bacterial promoters, E. coli strain MM294 (supra), Talmadge, K., et al., Gene (1980) 12:235; Messelson, M., et al., Nature (1968) 217:1110, was used as the host. For expression under control of the P_L N-RBS promoter, E. coli strain K12 MC1000 lambda lysogen, $N_7N_{53}cI857SusP_{80}$, ATCC 39531 (hereinafter sometimes referred to as MC1000-39513 λDG95 or DG95) may be used as well as E. coli strain DG116 also an MM294 strain (λ CI857, bio T76, del HI); the bio T76 substitution deletes early λ function (N att) and the del HI deletion removes λ DN from cro through att (del cro-Bio$^+$ n$^-$). This strain is deposited in the assignees culture collection under accession number CMCC 2298. For expression using pUC vectors E. coli DG99 may be used. DG99 has been deposited in the ATCC under Accession No. 39766.

For M13 phage recombinants, E. coli strains susceptible to phage infection, such as E. coli K12 strain DG98 are employed. The DG98 strains has been deposited with ATCC July 13, 1984 and has accession number 1965.

C. Specific Description of the Invention

The Streptomyces secretion vectors of the present invention comprise a plasmid replicable in Streptomyces, comprising a high level promoter operably linked to a DNA sequence encoding a Streptomyces signal sequence, wherein the DNA sequence encoding the secretion signal sequence may be operably linked to a heterologous DNA or gene encoding a desired protein to produce the desired protein in operable linkage with the Streptomyces signal sequence, and recite the desired protein. In a preferred embodiment, the DNA sequence encoding the secretion signal sequence further includes a site for insertion of the heterologous gene in operable linkage with the secretion signal sequence. High level promoters useful in the secretion vector according to the invention may be defined as promoters that are functional in Streptomyces to produce, in general, a Streptomyces gene product. High level promoters of this type are exemplified by promoters involved in the production of Streptomyces antibiotics such

as aminoglycoside phosphotransferase (aph), neomycin phosphotransferase (neo) and kanamycin phosphotransferase (kan). In addition, the promoters of genes which produce resistance to antibiotics, are also known to be high level promoters. Such high level promoters are exemplified by the erythromycin resistance gene (ermE). Such high level promoters are suitable for use in conjunction with a Streptomyces signal sequence and control the expression thereof.

Also within the scope of the invention, are modifications to such high level promoters thay may render them either more convenient or more effective in controlling the initiation of transcription and expression of the genes under their control. In the present invention, the high level promoter may be modified. For example, the ermE promoter is known to be a high level promoter. The ermE promoter has been modified to produce a promoter designated ermE1a. The sequence of the ermE promoter has been determined by Bibb et al., Gene 38:215-226, amended by the publisher in Gene 41:E357-E368 (1986). The modification of the ermE promoter to produce ermE1a is carried out by site-specific mutagenesis to remove the trinucleotide "TGG" in the (-35 region) of the P1 promoter of the ermE. As is shown in Figure 1, ermE has a number of overlapping promoters. ermE1a has been shown by Bibb et al. to yield greater levels of production of homologous Streptomyces gene products. In the present invention he ermE1a promoter has been further modified at the 3' end thereof to permit ligation of the ermE1a promoter to suitable Streptomyces signal peptides to create one embodiment of the Streptomyces secretion vector according to the invention. The 3' end of the ermE1a promoter is modified by first creating an SphI restriction site at the end of the ermE1a sequence by site-specific mutagenesis, followed by the removal of single-stranded DNA and digestion with EcoRI. The purified modified ermE1a promoter may be conveniently ligated with a suitable signal sequence as is shown in greater detail hereinbelow.

The aph promoter as mentioned above, is also a high level Streptomyces promoter. The aph promoter may also be modified so that it may be conveniently ligated to a signal sequence to produce the Streptomyces secretion vector coding to the invention.

The promoter of the aph gene of S. fradiae has been sequenced; see Bibb et al., Molec. Gen. Genet. 199:26-36 (1985). The sequence of the aph promoter is shown in Figure 2 hereof. A shortened fragment of the aph promoter was obtained by digestion of the aph promoter with BssHII and NcoI endonucleases, followed by repair of the single-stranded termini with Klenow fragment and dNTPs.

It will be appreciated from the foregoing that it may be necessary to make certain modifications to the high level promoter used to control the expression of the Streptomyces signal sequence in the vectors according to the invention. Such modifications are considered to be within the scope of the invention.

The Streptomyces signal peptide sequences appropriate for use in the present invention are exemplified by the signal secretion sequence of the ORF438 protein and the α-amylase signal peptide of S. limosus. The invention however is not considered to be limited to these two secretion signal sequences. The secretion signal sequences of S. plicatus endoH, S. avidinii streptavidin and S. coelicolor agarase are also believed to be suitable for use in vectors according to the invention. The signal peptide sequences are also modified in order to provide sites for insertion of heterologous genes in operable linkage with the secretion signal sequence. In general, the DNA encoding the 3' end of the signal sequence and the sequence encoding the NH2 terminal of the mature homologous protein that is normally produced under the control of the secretion signal sequence, are site-specifically modified to provide a number of restriction sites for the insertion of the DNA encoding the desired heterologous gene. The insertion sites are produced as restriction sites which allow the ligation of the DNA sequence encoding the desired heterologous gene to produce a fusion of the protein to the signal sequence encoded thereby preferably at the -1 or +1 residues wherein the -1 residue is the last residue of the secretion signal sequence and the +1 residue is the first residue of the mature desired heterologous protein. The foregoing modifications to the promoter or the secretion signal sequence, may be carried out of course by a variety of means. Oligonucleotide primer directed site-specific mutagenesis may be used

to modify the promoter or the signal secretion sequence as is shown in detail hereinbelow. However, other means may be used to produce the desired DNA sequence for the promoter or the signal secretion sequence. Indeed, it is possible to produce modified promoters or secretion signal sequences by synthesizing oligonucleotides encoding the required sequence and preferred restriction sites through various automated DNA synthesizers. A variety of DNA sequences encoding heterologous proteins may be expressed using the prominent secretion signal sequences of the present invention. Exemplified hereinbelow are but two of the human proteins, lymphotoxin and tumor necrosis factor that may be expressed and secreted using the vectors according to the invention. In addition, the procaryotic protein alkaline phosphotase obtained from E. coli is also produced in Streptomyces and secreted therefrom using the secretion vectors according to the invention.

The following examples are intended by the inventor to be exemplary and non-limiting.

Example I

I. Determination of the nucleotide sequence of the promoter proximal region of the α-amylase gene of S. limosus and the signal peptide sequence of the pre α-amylase.

A. Materials

Plasmid pIJ702 is a 5.8-kb, high copy number Streptomyces cloning vector (I). It is derived from S. lividans strain 66, and carries the tsr (thiostrepton resistance) and the mel (tyrosinase) genes. See Katz et al., J. Gen. Microbiol. (supra). pIJ702 can be obtained by reversing the orientation of the tsr, mel Bcl fragment of pIJ702, ATCC #39378. In addition pIJ702 is available as ATCC #37287. See p. 251, ATCC Catalogue, 16th ed. (1985).

Plasmid pIJ702.No512 is the original Streptomyces limosus amylase clone as a 6-kb insert in pIJ702. The pIJ2921.12SL plasmid is a pUC18 derivative of pIJ702.No5SL containing a 2.25-kb subcloned amy fragment flanked by KpnI sites. pUC18 is described in Gene 26:101-106

(1983) and is the same as pUC19 except that the polylinker sequence is in the opposite orientation. pUC19 is commercially available. See New England Biolabs Catalogue. The host strain used is S. lividans TK24 (pro-2, str-6, SLP2⁻, SLP3⁻); see Hopwood et al., J. Gen. Microbiol. 129:2257-2269. Other vectors are those frequently used in E. coli.

B. Mapping of the insert in pIJ2921.12SL

The restriction map of the insert of pIJ2921.12SL is shown in Figure 3. The location of the transcription start site maps just inside the insert in pIJ2921.12SL. The 5' overlapping fragment (2.3-kb SstI-EcoRI) from pIJ702.No5SL that contained the amy promoter sequence, was also cloned and used for mapping and sequence analysis.

C. Nucleotide sequence

Figure 4 shows the nucleotide sequence of the entire amy gene and the deduced amino acid sequence. Fragments containing portions of the amy gene were subcloned and the sequence determined by the dideoxynucleotide chain termination method. (See Messing et al., Nuc. Acid. Res. 9:309 (1981).) Additional sequencing primers were made according to preliminary sequence information, and they were used to generate further sequence data. The sequence shown here has been determined from both strands.

D. Reconstruction and cloning of the amy gene into pIJ702

To clone the amy promoter and the coding sequence of the amy gene with convenient flanking restriction sites, the following digestions and ligations were performed:

1. The 74-bp ApaI-KpnI fragment of pIJ702.No5SL was inserted into the HincII-KpnI site of pUC19 plasmid. This introduced the SphI site 5' to the original ApaI site located next to the amy promoter. The resulting intermediate plasmid was designated pSYC1344.

2. The 2.06-kb KpnI-BssHII fragment of pIJ2129.12SL (the BssHII end was first made flush by E. coli DNA polymerase Klenow fragment repair reaction) was isolated. It was cloned into the KpnI-HinCII region in pUC19. This introduced the PstI site at the 3' end

of the amy sequence. The resulting intermediate plasmid was designated as pSYC1346.

3. The two amy segments in these two plasmids were excised by SphI-KpnI and KpnI-PstI, respectively. The SphI-KpnI and KpnI-PstI fragments were ligated to the large SphI-PstI fragment of pIJ702 and, resulted in plasmid pSYC1388 (see Figure 3C). Plasmid pSYC1388 has been deposited in the assignees culture collection as CMCC No. 3027, and has been deposited in the American Type Culture Collection on March 4, 1987.

E. N-terminal sequence of mature α-amylase

Plasmid pSYC1388 was introduced into S. lividans strain TK24 by transformation. Transformations are picked by the loss of the dark colored phenotype conferred by the mel gene. One transformant was picked and grown in R6 medium at 30°C for 3 days. Cells were removed from the culture by centrifugation followed by filtration through millipore membrane. The exoproteins in the filtrate were concentrated and dialyzed by dialfiltration using a PM30 membrane. The sample containing the Streptomyces exoproteins was further fractionated on a Sephacryl S-200 column using sodium acetate buffer, pH 5.5, 2 mM DTT and 0.1% SDS. The purified α-amylase protein was subjected to amino-terminal sequence using an ABI gas-phase sequencer, and revealed the following sequence:

Ala-Pro-Pro-Gly-Ala-Lys-Asn-Val-Thr-Ala-Val-Leu-Phe-Glu-Trp

This sequence is in total agreement with the deduced sequence shown in Figure 4 and establishes the processing site by the signal peptidase as shown in that figure.

The following features of the amy promoter result from the sequencing data.

The transcription start site is the G located at position 243. This suggested that the "-35" and the "-10" sequences of the amy promoter are TTGACC and TACGGT, respectively.

The likely translational start is at position 311. The ATG triplet is preceded by the putative S/D sequence GGAGG located upstream.

The signal peptide starts from met, and ends at the 28th residue (the Ala at "-1").

An open reading frame (ORF) is located at the 5' end of this fragment. It is followed by an inverted repeat resembling the rho-independent transcription terminator sequence.

II. The construction and use of Streptomyces secretion vectors

A. Introduction and Materials

Vectors with the coding sequences of Streptomyces signal peptides under the control of high level promoters were constructed for the production of secreted heterologous proteins as follows:

The amylase gene from S. limosus (amy[Sli]); in pIJ2921.12 SL and pIJ702 were described in the previous example. The ORF438 gene is located upstream of tyrosinase gene (mel, for melannin) on pIJ702. See Bovnan et al., Gene 37:101-110 (1985). Plasmid pIJ61 is a 14.8-kb low copy number Streptomyces cloning vector. It is derived from the SLP1.2 plasmid from S. lividans 66, and carried the tsr (thiostrepton-resistance) and aph (aminoglycoside phophotransferase) genes. See D. A. Hopwood, M. J. Bibb, K. F. Chater, T. Kiester, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf, 1985, Genetic Manipulation of Streptomyces, A Laboratory Manual (The John Innes Foundation).

pIJ449 is a 3.2-kb plasmid containing the promoter region of the ermE (erythromycin-resistance) gene in the polylinker of pUC18 derivative. See M. J. Bibb, G. R. Janssen, and J. M. Ward, 1985, Gene 38:215-226 (amended by publisher in 1986, Gene 41:E357-E368).

B. Cloning and modification of the signal peptide coding region of amy (Sli) to produce secretion sites for insertion of heterologous DNA

The promoter-proximal region of the amy (Sli) gene was sequenced as described above. The 129-bp EcoRI-SmaI fragment from pIJ2921.12SL was ligated into the EcoRI-SmaI sites of RF DNA of M13mp10. This fragment contains the coding sequence for the amylase signal peptide. The Ala residue at position 29 (counting from the

putative initiation Met) is the last residue in the signal peptide prior to the proamylase protein sequence and Ala-28 is the "-1" residue as shown in Figure 1. The oligodeoxyribonucleotide primer SC64 was designed, chemically synthesized and used to modify the sequence by site-directed mutagenesis. The primer has the sequence 5'-CCCGGGCGGCGCCGCAGCGGCGGG-3'). The resulting RF DNA (designated pSYC1301) of the M13 phage has a new sequence for NarI (which is also BanI and HaeII) at the end of the coding sequence for amino acid #29, but still encodes the same peptide sequence.

```
original sequence:  ....CCC,GCC,GCT,GCC,GCC,CCG,....
                                     *   *
pSYC1301            ....CCC,GCC,GCT,GCG,GCG,CCG,....

                        -Pro-Ala-Ala-Ala-Ala-Pro-
                             26  27  28  29      (position in precursor)
                             -3  -2  -1  +1 +2   (position relative to
                                                  cleavage site)
```

Digestion with BanI and NarI, or HaeII of this modified sequence permits fusions at the -1 or +1 residues, respectively. This allows convenient manipulation to create fusions to the amy signal sequence.

pSYC1301 has been deposited in the applicants culture collection as CMCC number 3035 and was deposited with the American Type Culture Collection on March 4, 1987. The production of pSYC1301 is shown in Figure 5.

C. The cloning and modification of the signal peptide coding sequence of the ORF438 gene to produce restriction sites for the insertion of heterologous DNA

The 438-codon open reading frame (ORF438) that is linked to the mel gene has been described by Bernan et al., Gene, supra. As indicated by the authors, the ORF438 gene could potentially encode a secreted protein because its amino-terminal sequence resembles a signal peptide sequence. The signal peptidase cleavage site of the putative secretory protein was speculated to be between Ala-30 and Asp-31 as shown in Figure 6.

The PvuI site (CGATCG) of the ORF438 putative signal sequence was used for fusion of heterologous coding sequences at the +2 residue. In addition, the oliodeoxyribonucleotide primer SC69 was used to introduce a NarI site at the putative signal peptidase cleavage site by site-specific mutagenesis using an M13 vector. The primer had the sequence: 5'-CGCCGCCCGGGCGCCCGATCGGGG-3'.

One procedure used was as follows and is shown schematically in Figure 7.

The 433-bp BclI-BglII fragment of PIJ702 that contains the entire putative signal peptide coding sequence of ORF438 was cloned into the BamHI site pUC13. Insertion in both orientations were obtained; the 433-bp BclI-BglII fragment included the HincII site just upstream from the putative Shine/Dalgarno sequence of ORF438:

pSYC1259 (also pSYC1260). Two plasmids were obtained as follows:

EcoRI-SstI...(Bam/Bgl)-HincII-ORF438--PvuI--(Bcl/Bam)...XbaI-SalI-PstI-HindIII;

pSYC1261 (also pSYC1262):
HindIII-Pst-SalI-XbaI...(Bam/Bgl)-HincII-ORF438--PvuI--(Bcl/Bam)...SstI-EcoRI.

The EcoRI-HindIII fragment from pSYC1261 was cloned into M13mp11 at the corresponding sites, and SC69 primer was used to generate a NarI site by primer directed site specific mutagenesis. The following changes resulted:

```
                                  -3  -2  -1  +1  +2  +3
                          -Pro-Ala-Ala-Arg-Ala-Asp-Asp-Arg-Gly-
ORF438:      ...,CCC,GCC,GCC,CGC,GCG,GAC,GAT,CGG,GGG,....
                                      *   **
pSYC1392:    ...,CCC,GCC,GCC,CGG,GCG,CCC,GAT,CGG,GGG,....
                                        -Pro-
```

This mutation created a NarI site and changed the Asp at "+1" to a Pro. The RF DNA of the M13 phage was designated pSYC1392.

D. Cloning and modification of promoters for linkage to DNA encloding signal sequences for the expression of fusion genes

The promoters for the aminoglycoside phosphotransferase (aph) gene from S. fradiae (Bibb et al., Mol. Gen. Genet. (supra) and for the erythromycin-resistance (ermE) gene from S. erythreus (Bibb et al., Gene, supra) have been sequenced, and the transcriptional initiation sites determined as is shown in Figures 1 and 2.

1. aph-amy fusion: The aph promoter sequence flanked by the BssHII and the NcoI sites was excised from plasmid pIJ61. The termini of the fragment were repaired with Klenow fragment of DNA polymerase of E. coli, and blunt end ligated into the SmaI site of plasmid pUC13. The resulting plasmid pSYC1152 contained the aph promoter sequence in the orientation that the EcoRI site in the pUC13 polylinker is located 5' to the aph promoter. The NcoI site originally derived from the aph gene was regenerated. The same fragment inserted in the opposite orientation was also obtained, and it was designated as pSYC1151. The production of pSYC1151 and pSYC1152 are shown schematically in Figure 7. Plasmids pSYC1151 and pSYC1152 was deposited in applicants culture collection as CMCC #3010 and 3011, respectively, and were deposited in the American Type Culture Collection on March 4, 1987. The polylinker regions contain the sequences as shown below:

pSYC1152: EcoRI-SstI-.....aph promoter.....NcoI-BamHI-XbaI-SalI-PstI-HindIII

pSYC1151: HindIII-PstI-SalI-XbaI-BamHI-....aph promoter...NcoI-SstI-EcoRI

The pSYC1152 DNA was cleaved with NcoI, repaired with Klenow fragment and then cleaved by HindIII. The large HindIII-blunt end fragment was purified. The pSYC1301 DNA was cleaved with EcoRI, the termini were made flush with S1 nuclease, and then digested with HindIII. The small resulting fragment that contained the modified amy (Sli) sequence derived from pSYC1301 was isolated and was then ligated to the pSYC1152-derived fragment described above. The resulting plasmid pSYC1309 contains the aph promoter and the coding sequence of amylase signal peptide in the following order:

pSYC1309:
EcoRI-SstI...aph promoter...amy (Sli) signal peptide...NarI...---
HindIII.

Plasmid 1309 was deposited in applicants culture collection as CMCC #3012 and was deposited in the ATCC on March 4, 1987. The plasmid pSYC1309 was confirmed by restriction endonuclease digestion. The production of pSYC1309 is shown schematically in Figure 5.

2. ermE promoter-amy fusion: In order to create this fusion, an SphI site was first created 3'- to the ermE promoter as follows. Plasmid pIJ449 was digested with EcoRI and HindIII. The 0.59-kb fragment that contained the ermE promoter fragment was cloned into the corresponding site in M13mp10. The 23-nucleotide oligodeoxyribonucleotide primer SC65 having the sequence 5'-CCGAACTGCGCATGCGCTGGATC-3' was synthesized on an automatic DNA synthesizer and was employed for primer directed site specific mutagenesis which generated an M13 derivative with RF DNA that was designated as pSYC1303. The following changes in the ermE promoter were made:

```
original ermE:      (-10)------      fMet Ser Ser Ser Asp
                    .....TAGGATCCAGCG GTG AGC AGT TCG GAC .....
pSYC1303:           .....TAGGATCCAGCGCATG CGC AGT TCG GAC .....
                                      -------
                                       SphI
```

pSYC1303 was deposited in applicants culture collection as CMCC #3036 and was deposited in the ATCC on March 4, 1987.

The pSYC1303 DNA was digested with SphI, the single-stranded 3' tetranucleotides were removed by the exonuclease action of DNA polymerase. This DNA was then digested with EcoRI and the ermE promoter fragment was purified. This ermE-derived fragment, and the modified amy fragment prepared and isolated from pSYC1301 by the procedure described in the preceding paragraph, were ligated to pUC18 DNA that has been digested with EcoRI and HindIII. The resulting plasmid, pSYC1311, had ermE and amy in the following order:

pSYC1311:
EcoRI-SstI...ermE promoter...amy signal peptide coding sequence...NarI-HindIII

The production of pSYC1311 is shown schematically in Figure 5. Plasmid pSYC1311 has been deposited in applicants culture collection as CMCC #3013 and was deposited in the ATCC on March 4, 1987.

3. aph-ORF438 fusion:

Plasmid pSYC1151 carrying the modified aph promoter was digested with NcoI, the termini repaired with Klenow fragment, and then digested with HindIII. The resulting 0.2-kb fragment containing the aph promoter was cloned into plasmid pSYC1260 between the HindIII-HincII region, and generated pSYC1289 and pSYC1290. The production of pSYC1259, pSYC1260, pSYC1261 and PSYC1262 from plasmid pUC13 and pIJ702 is shown in Figure 7. Similarly, the aph promoter fragment was excised from pSYC1152 by digestion with NcoI, the ends filled in, and digested again with XbaI. This fragment was cloned into pSYC1262 between the XbaI and HincII sites, resulting in pSYC1291 and pSYC1292.

pSYC1289/1290:
HindIII-PstI-SalI-XbaI-BamHI...aph promtoer...ORF438 signal peptide...Sst-EcoRI

pSYC1291/1292:
EcoRI-SstI...aph promoter...ORF438 signal peptide...XbaI-SalI-PstI-HindIII

The production of pSYC1289/1290 and pSYC1291/1292 is also shown in Figure 7. Plasmids pSYC1259, 1261, 1290 and 1292 were deposited in applicants' culture collection as CMCC #'s 3016, 3017, 3018 and 3019, respectively. Plasmids pSYC1261 and pSYC1290 were deposited with the ATCC on March 4, 1987.

E. Expression of heterologous genes

1. TNF

a. Construction of aph-amy-TNF fusion

This fusion gene was constructed by ligation of four fragments and is schematically shown in Figure 8: (1) the large pUC13

fragment generated from EcoRI and SstI double digestion, (2) the SstI-NarI fragment from pSYC1309, which contains the modified aph promoter, (3) the amino terminal coding portion of TNF gene from M13 RF DNA LL02-4, generated from partial digestion with HapII and complete digestion with BstEII. This 385-bp fragment has a 5'-HapII site that was created by site-directed mutagenesis and (4) the 490-bp BstEII-EcoRI fragment from pAW711D that contains the C-terminal coding region of TNF and the cry terminator.

The resulting intermediate plasmid pSYC1376 which was confirmed by fragment sizing of endonuclease digested plasmid contains the following sequences:

HindIII-(linkers)-SstI..aph promoter..amy signal peptide..TNF..BamHI..cry..EcoRI

b. Construction of an ermE-amy-TNF fusion

Similarly, the 0.4-kb SstI-NarI fragment from pSYC1311 that contains the ermE promoter was used, and the intermediate plasmid pSYC1377 was constructed by the four-fragment ligation method as described in the previous section for the construciton of pSYC1376 and as schematically shown in Figure 8. The aph promoter fragment was substituted by the ermE fragment in pSYC1377.

The junction sequence between amylase signal peptide and the TNF sequence in both pSYC1376 and pSYC1377 is shown below:

amy-derived	TNF-derived
....amy.....CCC,GCC,GCT,GCG,GCG,GTC,....TNF...	
-Pro-Ala-Ala-Ala-Ala-Val-	

Plasmids pSYC1376 and pSYC1377 were deposited in applicants' culture collection as CMCC #'s 3014 and 3015, respectively.

c. Construction of a replicable vector for expression and secretion of TNF in Streptomyces lividans:

The construction of the expression vectors is shown schematically in Figure 8. The fragments containing the aph-amy-TNF and the ermE-amy-TNF fusions were purified from plasmids pSYC1376 and pSYC1377, respectively. These plasmids were first digested with

EcoRI, the ends repaired, and then digested with SstI. The repaired-EcoRI-to-SStI fragments were purified. Plasmid pIJ702 was digested with PstI, the ends were made flush by the exonuclease activity of DNA polymerase, and then digested with SstI. The large pIJ702-derived fragment was then ligated separately to each of the two promoter-amy-TNF fragments. The resulting Streptomyces plasmids pSYC1412 and pSYC1414 carry the aph-amy-TNF and the ermE-amy-TNF fusions, respectively.

S. lividans TK24 carrying pSYC1412 and pSYC1414 were grown in R6 medium for 3 days. The supernatant from these cultures were assayed for TNF biological activity. 1.3×10^4 U/ml of TNF using mouse cell line L929 to assay TNF activity, as described in U.S. Patent No. 4,667,064 issued June 30, 1987 and assigned to the assignee of the present application, was detected in the culture containing pSYC1412 (aph-amy-TNF) and 3.9×10^4 U/ml from pSYC1414 (ermE-amy-TNF). Cell density of these cultures was not determined.

The secreted TNF protein produced by the pSYC1414-transformed S. lividans strain was also analyzed. The proteins from 1.1 ml of culture supernatant from the control culture (pIJ702 plasmid with inactivated mel gene) and from the test culture (pSYC1414 plasmid) were precipitated by 20% TCA. Samples were then fractionated on a 12.5% SDS-PAGE by electrophoresis. The result is shown in Figure 9. The secreted TNF is estimated to be in the range of 1 mg/l, which is consistent with the bioassay result. This material was purified and its amino-terminal sequence determined. The junction sequence between the amylase signal peptide and TNF is:

```
  -4  -3  -2  -1                          (position in amylase)
-Pro-Ala-Ala-Ala-Ala-Val-Arg-Ser-Ser-Ser-
                        "+1"              ***
                      in TNF              (observed start of
                                           secreted TNF)
```

The observed start is the Ser as marked above. Four residues of the mature TNF was lost, presumably due to peptidase activities in the culture supernatant. As can be seen in Figure 9, there is a minor band above the major TNF protein band that may

represent the primary secreted TNF protein, or an intermediate degradation product. However, TNF lacking the first four amino acid residues is biologically active.

Plasmids pSYC1412 and pSYC1414 have been deposited in applicants depository as CMCC #'s 3028 and 3029, respectively and were deposited in the ATCC on March 4, 1987.

2. Secretion of alkaline phosphatase (AP) of E. coli in Streptomyces

The production of the secretion for E. coli alkaline phosphatase in Streptomyces is shown schematically in Figure 10. The E. coli plasmid pSYC1204 contains the phoA gene with the cry gene terminator at the 3'-end of the phoA gene. A plasmid having the phoA gene and cry terminator may be constructed in accordance with the disclosure of European Patent Application No. 86302201.8 filed March 25, 1986 An NruI sequence may also be introduced into the phoA leader. Alternatively, the DNA sequence of the modified phoA leader and cry terminator may be synthesized as oligodeoxyribonucleotides. pSYC1204 has the following organization:

...HindIII....phoA promoter..phoA signal peptide..NruI..AP....BamHI.. cry terminator..EcoRI....

An NruI site was introduced at the coding sequence for the "-1" residue of the pre-AP protein. The plasmid pSYC1204 was digested with HindIII and NruI, and the large fragment, containing the coding sequence for mature AP, purified. Plasmid pSYC1259 was digested with PvuI (which is located at the sequence corresponding to the second codon of the putative mature ORF438 protein), the single-stranded dinucleotide was trimmed away with the exonuclease activity of DNA polymerase, and then digested further with HindIII. The small fragment carrying the ORF438 promoter and putative signal peptide coding sequence was ligated to the pSYC1204-derived large fragment under blunt end conditions. This resulted in the plasmid pSYC1267. The ORF-phoA fusion sequence was excised from pSYC1267 by BamHI and PstI digestion, and was cloned into the PstI-BglII sites in pIJ702 which generated the Streptomyces plasmid pSYC1283. S. lividans

carrying pSYC1283 was grown in R6 medium for 3 days and produced low level of AP in the culture supernatant as determined by a conventional spectrophotometric assay using p-Nitrophenyl phosphate (disodium) as described in J. Bacteriol., 152:692 (1982). The AP protein could not be identified from the total proteins in 1 ml culture by coomasie blue staining of the gel-banded proteins.

A plasmid was also constructed that expressed phoA under the control of the aph promoter. Plasmid pSYC1283 was partially digested with SphI (one of the SphI sites on the plasmid is located in the ORF438 sequence corresponding to the translation initiation site), and then digested with PstI which is located upstream of the phoA sequence. The fragment that had the ORF438 promoter and Shine/Dalgarno sequence deleted was purified and ligated to the PstI-SphI fragment of pSYC1290 that carried the aph promtoer and the ORF438 Shine/Dalgarno sequence. This generated the Streptomyces plasmid pSYC1328.

The supernatant of a 3-day culture of S. lividans carrying pSYC1328 grown in R6 medium was concentrated by diafiltration using a PM30 membrane. The exoproteins were analyzed by SDS-PAGE as shown in Figure 11. The secreted AP represents a major species of exoprotein synthesized. Surprisingly, the level is estimated to be around 5 mg/l. This concentrated exoprotein preparation was then fractionated in the presence of 0.1% SDS through a S-200 Sephacryl column. The secreted AP was then subjected to amino terminal sequence analysis. The ORF438-phoA fusion has the following junction sequence:

```
position in ORF438:  -4  -3  -2  -1  +1  +2
position in AP:                          +1  +2  +3  +4
                    -Ala-Ala-Arg-Ala-Asp-Ala-Arg-Thr-Pro-Glu-
```

The sequencing data showed that about 1/3 of the purified AP has the Ala-Asp-Ala-Arg-Thr-Pro-....sequence, about 2/3 started with the authentic mature AP sequence Arg-Thr-Pro-Glu..... It is unclear whether the ORF-AP "pre" protein was processed at two different sites that yielded these two species of proteins, or that there was only a single signal peptidase processing site, but some of the protein was further degraded. In fact, heterogeneity at the amino terminus of

secreted proteins is not uncommon in many systems; one example in Streptomyces is the endo-H protein from S. plicatus. See Robbins et al., J. of Biol. Chem. 259:7572-7581 (1984). The result is also surprising in that it suggests that the Aph promoter functions with great efficiency in the presence of a downstream Shine/Dalgarno sequence.

Example 2

Synthesis and Molecular Cloning of the Human Lymphotoxin Gene and its Expression in Streptomyces lividans

In this example, a synthetic ltx gene was assembled in E. coli and expressed in Streptomyces lividans. Secreted lymphotoxin was produced as a major protein in the medium at the level of several microgram per milliliter in shake-flask culture.

To assemble the synthetic ltx gene, oligodeoxyribonucleotide fragments shown in Figure 12 by the double lines were synthesized using an automated DNA synthesizer and were individually kinased. Synthetic fragments, each in 12 µl volume at 17 pmole/µl, were combined in pairs. To this, equal volume (24 (µl) of 2 x annealing buffer (200 mM NaCl, 40 mM $MgCl_2$, 40 mM Tris-HCl, pH 7.9, and 40 mM ß-mercaptoethanol) was added. This mixture was then heated to 70°C and gradually cooled to 10°C over several hours.

In greater detail, the ltx gene sequence has been reported previously (P. W. Gray et al., Nature 312:721-724 (1984)). A small number of long synthetic fragments were annealed and formed a partially double-stranded form. The gaps were then filled in by E. coli DNA polymerase in vitro to generate completely double-stranded DNA for cloning.

The synthetic ltx gene is shown in Figure 12. It has the following features:

1. It encodes the mature ltx proteion of 171 amino acids.

2. It has a total of 1074 (2 x 537) nucleotides; of which 642 (60%) were made by chemical synthesis in ten fragments, the rest were by enzymatic synthesis.

3. Unique restriction sites, which facilitated subsequent construction for cloning and expression, were introduced at the termini and the internal regions.

B. Assembly of the Synthetic Fragments

The strategy was to anneal the fragments in vitro to form full-length (or greater size) molecules with gaps, and then to fill in the gapped regions with DNA polymerase, and to seal the joints with DNA ligase. The fragments were first annealed in pairs so that each duplex contained 5'-protruding ends to allow DNA polymerase-directed synthesis. The five pairs consisted of (a) fragments 1+2, (b) 3+4, (c) 5+6, (d) 7+8, and (e) 9+10.

Three pairs of the annealed fragments (c), (d), and (e) were then combined (total of 144 µl), and 30 µl of 2.5 mM dXTP, 3 µl of 10 mM ATP, 2 µl of DNA polymerase Klenow enzyme, 2 µl of T4 DNA ligase were added. The reaction was carried out at 4°C for 20 minutes followed by incubation at 20°C for 60 minutes. The reaction was then stopped and the DNA extracted. A fully double-stranded DNA was synthesized. The DNA fragments were digested with HindIII and BamHI, purified from gel, and cloned into the corresponding sites of pUC13 plasmid vector. From plasmid screening of 32 candidates, 24 correct clones were identified. Plasmid pSYC1378 was a representative in this group.

A similar approach to assemble pairs (a), (b), and (c) into a single long fragment was unsuccessful. However, the double-stranded DNA that was produced was digested with HindIII and the resulting fragments were cloned into a HindIII/SmaI doubly digested pUC13 vector DNA. Although none of the candidates screened had the structure completely correct, clones with partially correct structure by restriction endonuclease mapping were identified. Among the double stranded fragments were found those that carry the correct sequence in the regions covering the EcoRI-NsiI, the NsiI-KpnI, and the KpnI-HindIII sequences. These three fragments were cloned along with the HindIII-BamHI fragment from pSYC1378 into the EcoRI-HindIII sites of

pUC18. The derivative that contained the ltx sequence was designated pSYC1408.

C. Expression in Streptomyces lividans of the ltx gene

Two plasmids were constructed to express the ltx gene. The amy-ltx fusion plasmid consists of four parts generated by ligation of the following fragments:

1. the ltx fragment - The ltx gene in pSYC1408 was excised with HinPI (an isoschisomer of HhaI) and BamHI (the HinPI-generated termini can be ligated to NarI-generated fragments).

2. the B. thuringiensis retroregulator (the cry terminator) flanked by BamHI and EcoRI from plasmid pLW1t-10. The cry terminator may also be obtained in the same BamHI-EcoRI fragment from pHCW701, ATCC #39757.

3. the ermE-amy fusion fragment (ermE promoter and amylase signal peptide coding sequence), described above, flanked by SstI-NarI from pSYC1311.

4. the large pUC13-derived fragment generated by SstI and EcoRI.

The resulting plasmid was designated pSYC1443. The ermE-amy-ltx-cry fusion was excised from pSYC1443 with EcoRI. The termini were repaired with Klenow fragment and the fragment was then digested with SstI. It was ligated into the Streptomyces plasmid vector pIJ702 under blunt end conditions between the PstI site that had been trimmed with S1 nuclease to form flush end and the SstI site. The resulting plasmid, pSYC1481, was introduced into Streptomyces lividans strain TK24 by protoplast transformation for expression studies.

Similarly, the ltx gene was fused to the signal sequence of ORF438. The modified ORF438 sequence with a NarI site introduced at the end of the signal peptide coding sequence as described above was obtained from pSYC1292 by HincII and XbaI digestion. The ermE promoter fragment with the added SphI site was excised from pSYC1303 with SphI; the termini were then trimmed with the exonuclease activity

of DNA polymerase and digested with EcoRI. These two fragments were then ligated into the EcoRI-XbaI region of pUC18. The resulting plasmid, pSYC1449 (and also pSYC1448), was employed for expression vector construction.

The 0.4-kb SstI-NarI fragment containing the ermE-ORF fusion (ermE promoter and the signal peptide coding sequence of ORF438) was purified after partial digestion with NarI and cleavage with SstI of the pSYC1449 DNA. It was ligated with ltx from pSYC1408, the cry gene retroregulator from pLW1t-10, and the pUC13 plasmid fragments, as outlined above for the construction of pSYC1443. The resulting plasmid, pSYC1450, contained the promoter of ermE followed by sequences derived from ORF, ltx and cry retroregulator. This region was then excised with EcoRI; the termini were repaired with Klenow fragment and then digested with SstI, and cloned into the PstI site, which was trimmed with S1 nuclease, and SstI site of the Streptomyces plasmid pIJ702 to generate plasmids pSYC1483 and pSYC1481. Plasmid pSYC1483 has been deposited in applicants' culture collection as CMCC #3031 and was deposited in the ATCC on March 4, 1987.

Streptomyces lividans strain TK24 transformed with either pSYC1481 or pSYC1483 and grown in shake flasks for three days in R6 medium and the extracellular lymphotoxin activity was determined using the assay mentioned above for TNF activity. TK24 (pSYC1481) produced about 500 U/ml, wherease TK24 (pSYC1483) produced 4×10^3 U/ml. The level of TK24 (pSYC1483) reached 4×10^4 U/ml on day 5 and 1.2×10^5 U/ml on day 7.

The extracellular proteins from TK24 (pSYC1483) were analyzed by SDS-PAGE as shown in Figure 13. The apparent 16-kd protein is slightly smaller than the expected size for mature ltx or 171 residue. The intensity of this putative ltx protein band is consistant with the activity assay data, and it represents about 2-4 µg/ml of lymphotoxin in the medium on day 7.

Example 3

Modification of promoter-signal sequence construction for increased secreted expression

The chimeric amy signal sequence-TNF gene under the control of the aph or the ermE promoters expressed TNF at relatively lower levels in Streptomyces lividans in the example above. A similar construction using the ORF438 signal peptide fragment that was used in the expression of lymphotoxin was made.

The construction of ORF signal peptide-TNF fusion gene is shown in Figure 14. In plasmid pSYC1449, the 0.28-kb ermE promoter fragment is located upstream of the 0.12-kb fragment that contains the ORF438 ribosome binding site and signal peptide coding sequences. By a four-fragment ligation procedure, the ermE prompter, the ORF438 ribosome binding site, the ORF438 signal peptide coding sequence, the TNF coding sequence, and the B.t. retroregulator were linked to a pUC13 vector. pSYC1449 was partially digested with NarI, and digested to completion with SstI. Plasmids pUC13, M13pAW711 and pAW711D were digested as described above in connection with the production of plasmid pSYC1414. The required fragments could also be obtained from pSYE1414. The large EcoRI-SstI pUC13 fragment, 385-bp HpaII-BstEII fragment of M13pAW711, 490-bp BstII-EcoRI fragment of pAw711D were combined with and ligated to the NarI-SstI fragment of pSYC1449 under sticky end conditions. A plasmid having the fragments in proper order as determined by restriction endonculease digestion and mapping was designated pSYC1477.

The ermE-ORF-TNF fusion gene from pSYC1447 was excised by EcoRI, followed by repair with Klenow fragment to generate flush ends and digested with SstI. The SstI-blunt fragment was inserted into the Streptomyces vector pIJ702 between the PstI (blunt-ended) and the EcoRI sites. The plasmid generated by this procedure was designated pSYC1493. pSYC1493 has been deposited in the applicants' culture collection as CMCC #3032 and was deposited in the ATCC on March 4, 1987.

A modified ermE promoter, ermE1a, was also used in the analogous construction. This promoter differs from its parent by the deletion of the TGG triplet near the "-35" region of the ermE promoter (see Figure 1).

Figure 15 outlines the construction of the plasmid pSYC1504, which is similar in pSYC1493 except that the ermE promoter in pSYC1493 was replaced by the ermE1a promoter carried on plasmid pIJ4065. The modification of the ermE promoter to produce ermE1a promoter can also be carried out using site-specific mutagenesis of the small BamHI to BamHI fragment of pSYC1472 to yield the same result.

S. lividans TK24 transformants carrying pSYC1493 or pSYC1504 were grown in shake flasks in R6 (trypton soy broth) medium for 3 days. TNF proteins in the culture supernatants and cell pellets were then analyzed.

Table 2 shows the TNF activity assay results of various TK24 transformants. The TK24 (pSYC1493) strain produced about 7-fold more TNF than TK24 (pSYC1414). These two strains differ in two aspects - the ribosome binding site and the signal peptide coding sequences. The ORF signal peptide, together with its ribosome binding site, is a better combination then the cognate sequences from the amy (Sli) gene.

Strain TK24 (pSYC1504) showed a further 2-fold increase in TNF expression than TK24 (pSYC1493).

TABLE II

Secreted TNF form TK transformants

Plasmid	Promoter	S/D + Sig. Pep.	Secreted TNF U/ml; (µg/ml)*	Junction Sequence (Signal peptide--TNF)
pSYC1412	aph	amy	1.3×10^4	...PAAA-AVRSSS...
pSYC1414	ermE	amy	4.0×10^4	...PAAA-AVRSSS...
pSYC1493	ermE	ORF438	2.7×10^5 (~10)	...PAARA-VRSSS...
pSYC1504	ermE1a	ORF438	5.5×10^5 (~20)	...PAARA-VRSS...

*Specific activity = $2\text{-}3 \times 10^7$ U/mg

The total cellular proteins and the secreted proteins from these strains were analyzed. As shown in Figure 16, which shows 12.5% SDS-page preparation of culture supernatants, TNF is the most abundant protein species in the three-day culture medium supernatants from the two TNF-producing strains. From the stained gel, it is estimated that TK24 (pSYC1504) produced about 0.5 µg of TNF in 25 µl of culture-equivalent to 20 mg/l. This estimate matches the activity assay result shown in Table II.

Western analysis of cellular proteins (Figure 17) showed that there were three TNF-related protein species in the cell, one in the culture supernatant. The lowest band corresponds to the "-4" form of TNF. The highest band likely represents the full-length "pre-TNF" protein with the 30-residue signal peptide sequence still linked to TNF. The middle band located just above the "-4" TNF protein is believed to represent the processed TNF generated from pre-TNF by signal peptidase cleavage.

To the extent that full-length mature TNF is represented by the middle band, then the amino-terminal four residues were removed from this protein before it was released into the medium. The detection of these two forms of matured TNF proteins in the cell indicates that processed TNF molecules are initially trapped in the cell, most likely outside the cell membrane but inside the cell wall.

We propose the following post-translational processing events that lead to the secretion of TNF:

Pre-TNF having the recombinant signal sequence is produced inside the cytoplasmic membrane. A signal peptidase active on the membrane processes the pre-TNF and secretes a full-length mature TNF associated with the cell envelope, but outside of the cytoplasmic membrane. An aminopeptidase associated with the cell envelope further processes the full length mature TNF to produce the "-4" form of TNF which is released into the medium as secreted "-4" TNF.

As will be apparent to those skilled in the art, the secretion vectors according to the invention are useful in the secreted production in Streptomycetes of heterologous proteins from

various sources, including mammalian protein. The secretion vectors with heterologous DNA other than that specifically exemplified are, of course, considered to be within the scope of the invention.

WHAT IS CLAIMED IS:

1. A recombinant vector suitable for secreting heterologous proteins in streptomycetes comprising a plasmid replicable in streptomycetes, said plasmid comprising a promoter functional in Streptomyces, and a DNA sequence encoding a secretion signal sequence that is operable in streptomycetes and may be operably linked at a site for insertion of a heterologous gene under the control of said promoter.

2. The recombinant vector of claim 1 further including a heterologous gene in operable linkage with the secretion signal sequence.

3. The recombinant vector of claim 1 wherein the promoter is selected from the group consisting of the aph, ermE promoter, ermE1A and modified ermE1a.

4. The recombinant vector of claim 1 wherein the signal sequence is a streptomycetes signal sequence.

5. The recombinant vector of claim 4 wherein the signal sequence is selected from the group consisting of the amy signal sequence and the ORF438 signal sequence.

6. The recombinant vector of claim 1 comprising the aph promoter and the secretion signal sequence selected from the group consisting of the amy secretion signal sequence, and the ORF438 secretion signal sequence.

7. The recombinant vector of claim 6 wherein the ORF438 secretion signal sequence and the aph promoter are separated by an intervening DNA sequence comprising a Shine/Dalgarno sequence.

8. The vector of claim 1 wherein the promoter is ermE-SphI and the secretion signal sequence is selected from the group consisting of the amy secretion signal sequence ORF438.

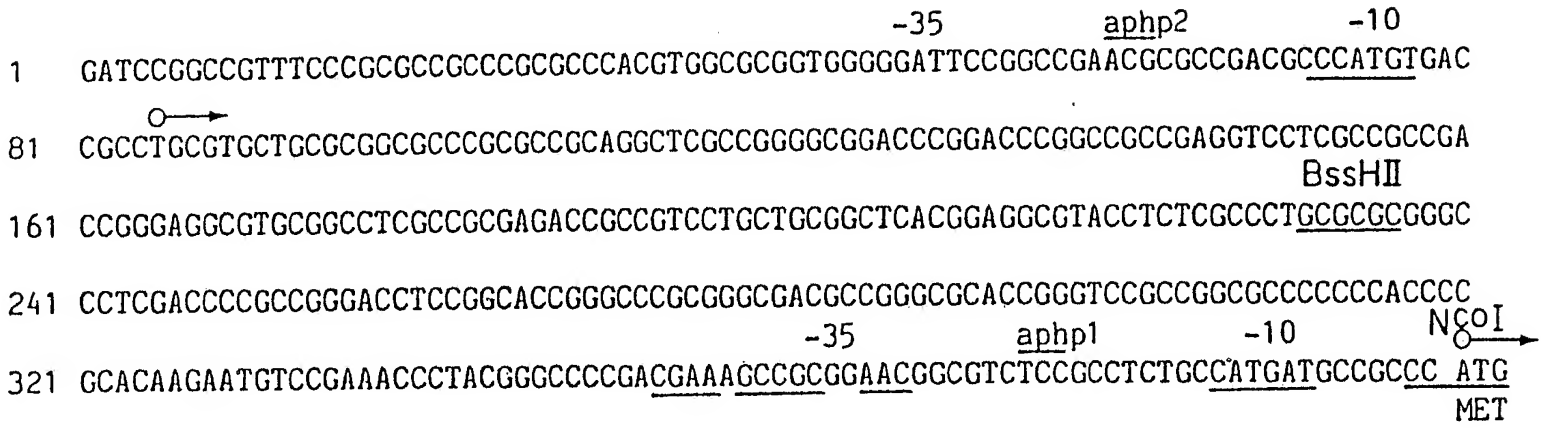
9. The vector of claim 1 wherein the heterologous gene encodes a protein of a eukaryotic or prokaryotic protein.

10. A host streptomycetes cell transformed with the vector of claims 1, 2, 3, 4, 5, 6, 7, 8 or 9.

11. The host cell of claim 15 wherein said streptomycetes are selected from the group consisting of S. lividans, S. coelicolor, S. griseus, S. parvulus, S. albus, S. vinaceus, S. acrimycini, S. calvuligerus, S. rubiginosis, S. azureus, S. glaucescens, S. rimosus, S. limosus, S. kanamyceticus, S. violaceotuber, and any streptomycetes related host in which the replicon of PIJ702 is functional.

Promoter region of ermE gene of Streptomyces erythreus

FIG._1.



Promoter region of the aph gene of Streptomyces fradiae

FIG. 2.

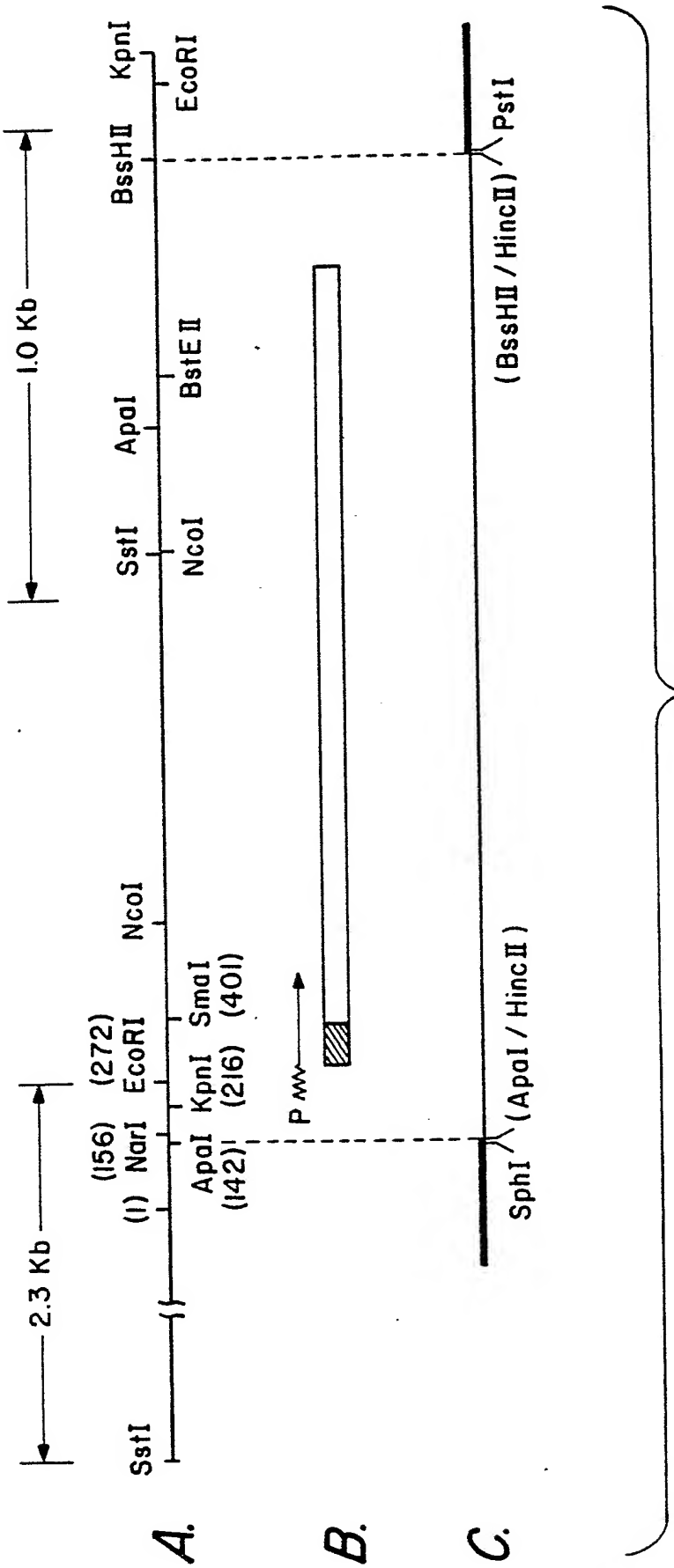

1.0 Kb
2.3 Kb
SstI
(1) NarI (156)
(142)!! ApaI
KpnI (216)
EcoRI (272)
SmaI (401)
NcoI
P
SstI
NcoI
ApaI
BstE II
BssHII
KpnI
EcoRI
SphI
(ApaI / HincII)
(BssHII / HincII)
PstI
A.
B.
C.

FIG._3.

FIG._4.

```
   1 CCAGACCGTGCAGCGCACCGAGTTCGTCTTCCAGCCGGAGCTGGTGGTGCGCGGCTCCAC
     GlnThrValGlnArgThrGluPheValPheGlnProGluLeuValValArgGlySerThr

                                                >>>>..>>> >>..<< <<<..<<<<
  61 CGCCCAGTGGGTGCCCGGCGGCTGAGGGGATGTACGGGGCGCTTCGTCCGGAGCCGCCCG
     AlaGlnTrpValProGlyGly...

 121 CAATCTCTTGCAGAGGCTTGCGGGCCCTGCTGCCCGGCGCCCAACCCCCTTGATGTCCGT

 181 GGTGTTTCGGGTTCGTGTCCAAAGGGTTGACCGCGGGTACCGCTCGCTCTACGGTCTGCT

       !(+1)                                                    =
 241 TCGCGACGCTCCTTGCAGTTTTGCTGCAAGAGAATTCAGCCCTCCGCCCCCCCCGATCAG

     ====
 301 GAGGCACCACATGGCCCGCAGACTCGCCACCGCGTCCCTAGCCGTGCTGGCGGCGGCCGC
               METAlaArgArgLeuAlaThrAlaSerLeuAlaValLeuAlaAlaAlaAla
               -28      -25             -20             -15

 361 CACCGCCCTCACCGCGCCCACACCCGCCGCTGCCGCCCCGCCCGGGGCGAAGGACGTCAC
     ThrAlaLeuThrAlaProThrProAlaAlaAlaAlaProProGlyAlaLysAspValThr  9
        -10             -5          -1 ▲1

 421 CGCCGTCCTCTTCGAGTGGAAGTTCGCCTCCGTAGCCCGCGCCTGCACCGACAGCCTCGG
     AlaValLeuPheGluTrpLysPheAlaSerValAlaArgAlaCysThrAspSerLeuGly 29

 481 CCCGGCCGGCTACGGATACGTCCAGGTCTCGCCGCCCCAGGAGCACATCCAGGGCAGCCA
     ProAlaGlyTyrGlyTyrValGlnValSerProProGlnGluHisIleGlnGlySerGln 49

 541 GTGGTGGACCTCCTACCAGCCCGTCAGCTACAAGATCGCCGGACGGCTCGGCGACCGCGC
     TrpTrpThrSerTyrGlnProValSerTyrLysIleAlaGlyArgLeuGlyAspArgAla 69

 601 CGCCTTCAAGTCCATGGTCGACACCTGCCACGCGGCCGGCGTCAAGGTCGTCGCCGACTC
     AlaPheLysSerMETValAspThrCysHisAlaAlaGlyValLysValValAlaAspSer 89
```

FIG._4-1.

661 GGTCATCAACCACATGGCCGCGGGTTCCGGCACCGGCACCGGCGGCAGCGCGT
ValIleAsnHisMETAlaAlaGlySerGlyThrGlyThrGlyGlySerAlaT

721 GTACGACTACCCGGGCATCTGGTCCGGCGCCGACATGGACGACTGCCGCAGCGAGATCAA
TyrAspTyrProGlyIleTrpSerGlyAlaAspMETAspAspCysArgSerGluIleAsn129

781 CGACTACGGCAACCGCGCCAACGTCCAGAACTGCGAACTGGTCGGCCTCGCCGACCTCGA
AspTyrGlyAsnArgAlaAsnValGlnAsnCysGluLeuValGlyLeuAlaAspLeuAsp149

841 CACCGGTGAGTCGTACGTCCGCGACCGCATCGCCGCCTACCTCAACGACCTGCTCTCGCT
ThrGlyGluSerTyrValArgAspArgIleAlaAlaTyrLeuAsnAspLeuLeuSerLeu169

901 CGGTGTGGACGGCTTCCGCATCGACGCCGCCAAGCACATGCCCGCCGCCGACCTCACCGC
GlyValAspGlyPheArgIleAspAlaAlaLysHisMETProAlaAlaAspLeuThrAla189

961 CATCAAGGCCAAGGTCGGCAACGGGAGCACGTACTGGAAGCAGGAGGCCATCCACGGCGC
IleLysAlaLysValGlyAsnGlySerThrTyrTrpLysGlnGluAlaIleHisGlyAla209

1021 GGGCGAGGCCGTCCAGCCCAGCGAGTACCTCGGCACCGGCGACGTCCAGGAGTTCCGCTA
GlyGluAlaValGlnProSerGluTyrLeuGlyThrGlyAspValGlnGluPheArgTyr229

1081 CGCCCGCGACCTCAAGCGGGTCTTCCAGAACGAGAACCTCGCCCACCTGAAGAACTTCGG
AlaArgAspLeuLysArgValPheGlnAsnGluAsnLeuAlaHisLeuLysAsnPheGly249

1141 CGAGGACTGGGGCTACATGGCGAGCGGCAAGTCCGCCGTCTTCGTCGACAACCACGACAC
GluAspTrpGlyTyrMETAlaSerGlyLysSerAlaValPheValAspAsnHisAspThr269

1201 CGAGCGGGGCGGCGACACCCTCAACTACAAGAACGGCTCCGCCTACACCCTCGCCGGCGT
GluArgGlyGlyAspThrLeuAsnTyrLysAsnGlySerAlaTyrThrLeuAlaGlyVal289

1261 CTTCATGCTGGCCTGGCCCTACGGCTCCCCGGACGTCCACTCCGGCTACGAGTTCACCGA
PheMETLeuAlaTrpProTyrGlySerProAspValHisSerGlyTyrGluPheThrAsp309

1321 CCACGACGCCGGCCCGCCCAACGGCGGCACCGTCAACGCCTGCTACAGCGACGGCTGGAA
HisAspAlaGlyProProAsnGlyGlyThrValAsnAlaCysTyrSerAspGlyTrpLys329

1381 GTGCCAGCACGCCTGGCCCGAGCTCTCCTCCATGGTCGGCCTGCGCAACACCGCCTCCGG
CysGlnHisAlaTrpProGluLeuSerSerMETValGlyLeuArgAsnThrAlaSerGly349

1441 GCAGCCCGTCACCAACTGGTGGGACAACGGCGGCGACCAGATCGCCTTCGGCCGCGGCGA
GlnProValThrAsnTrpTrpAspAsnGlyGlyAspGlnIleAlaPheGlyArgGlyAsp369

1501 CAAGGCGTACGTCGCCATCAACCACGAGGGCTCCGCGCTGAACCGCACCTTCCAGAGCGG
LysAlaTyrValAlaIleAsnHisGluGlySerAlaLeuAsnArgThrPheGlnSerGly389

1561 CCTGCCCGGCGGCGCCTACTGCGACGTCCAGAGCGGCAGGTCCGTCACGGTCGGCTCCGA
LeuProGlyGlyAlaTyrCysAspValGlnSerGlyArgSerValThrValGlySerAsp409

1621 CGGCACCTTCACCGCCACCGTCGCCGCCGGCACCGCCCTGGCCCTGCACACCGGGGCCCG
GlyThrPheThrAlaThrValAlaAlaGlyThrAlaLeuAlaLeuHisThrGlyAlaArg429

1681 TACCTGCTCCGGCGGCGGAACCGGCCCCGGCACCGGGCAGACCTCCGCCTCCTTCCACGT
ThrCysSerGlyGlyGlyThrGlyProGlyThrGlyGlnThrSerAlaSerPheHisVal449

FIG._4-2.

```
1741 CAACGCCACCACCGCCTGGGGCGAGAACATCTACGTCACCGGTGACCAGGCCGCCCTCGG
      AsnAlaThrThrAlaTrpGlyGluAsnIleTyrValThrGlyAspGlnAlaAlaLeuGly469

1801 CAACTGGGACCCGGCCCGCGCCCTCAAGCTCGACCCGGCCGCCTACCCGGTGTGGAAGCT
      AsnTrpAspProAlaArgAlaLeuLysLeuAspProAlaAlaTyrProValTrpLysLeu489

1861 CGACGTGCCGCTGGCCGCCGGAACCCCCTTCCAGTACAAGTACCTGCGCAAGGACGCCGC
      AspValProLeuAlaAlaGlyThrProPheGlnTyrLysTyrLeuArgLysAspAlaAla509

1921 GGGGAAGGCCGTCTGGGAGTCCGGCGCCAACCGCACGGCGACCGTCGGCACCACCGGCGC
      GlyLysAlaValTrpGluSerGlyAlaAsnArgThrAlaThrValGlyThrThrGlyAla529

                                                    >>>>>>>>>>.>>>>..
1981 CCTCACCCTCAACGACACCTGGCGCGGCTGACCCCGCCCCCATCGGCCCGGCGCCCGCCT
      LeuThrLeuAsnAspThrTrpArgGly...
                                538

     ....<<<<.<<<<<<<<<<
2041 CCCCGCGGACGCCGGGCCGCTGCCGCACCCGCCCCGCGAACCCCGCCCGGCGGCCCCCGC

2101 CCCCGAAGGAGACCCGTCCCCGTGCGACGGACCACCATCCTCGCGGTGAGTCTCGGCCTG

2161 TGCGCCGCCCTCACCGCCACCCTCCCGGCCACCGCCGACACCCCCGACGCCCCGGCCGCG

2221 CGGGCCACCGCCGCCGAGGACACCGCCGAGGCCGTCTGGCTCGACGCCCGCACCGTCGCC

2281 TGGCCGCGCGC
```

FIG._4-3.

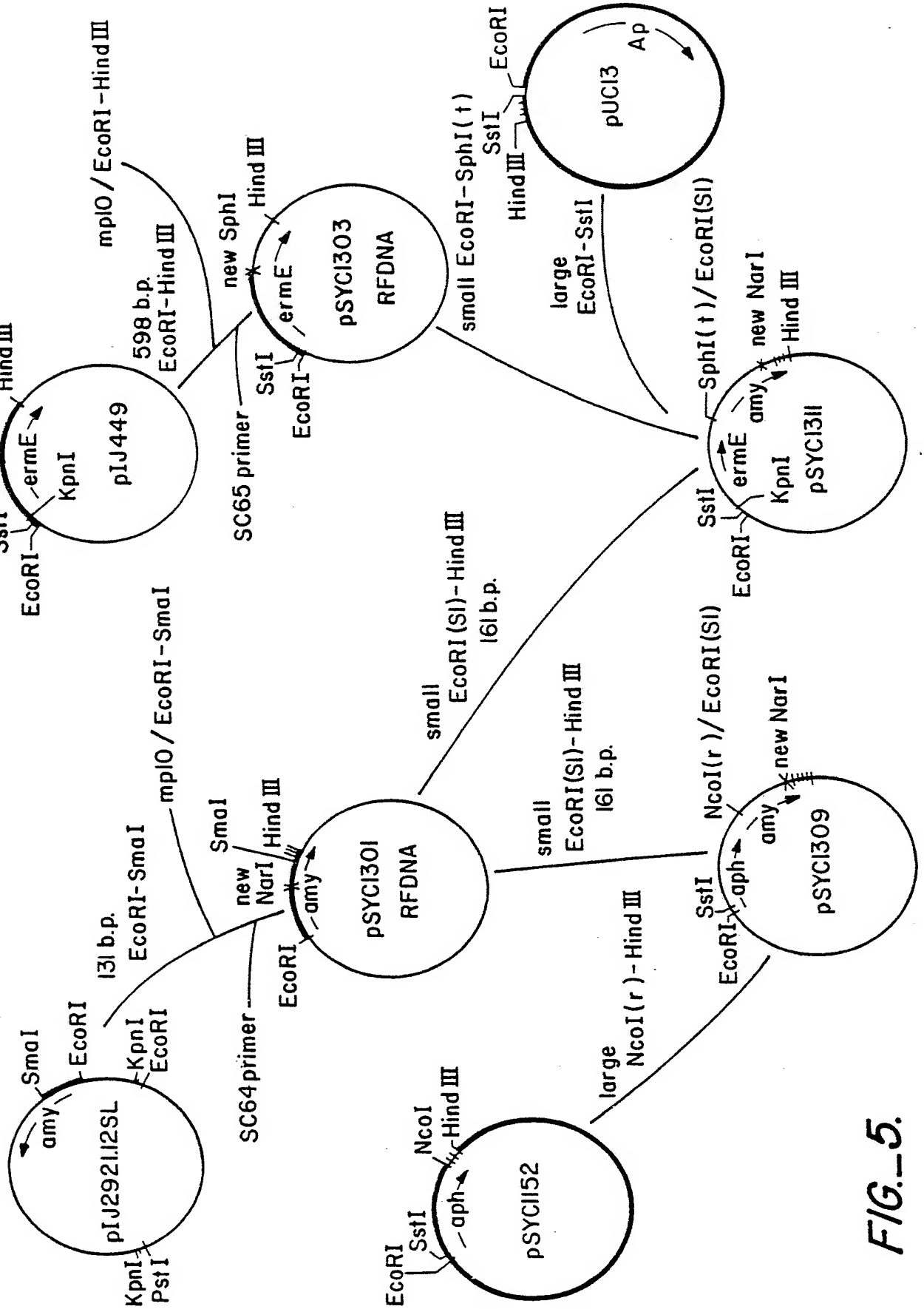

pIJ2921.12SL
amy
SmaI
EcoRI
KpnI
EcoRI
KpnI
PstI
131 b.p.
EcoRI-SmaI
mp10/EcoRI-SmaI
SC64 primer
pSYC1152
aph
NcoI
HindIII
SstI
EcoRI
large
NcoI(r)-HindIII
pSYC1301
RFDNA
new
NarI
SmaI
HindIII
amy
EcoRI
small
EcoRI(SI)-HindIII
161 b.p.
pSYC1309
aph
NcoI(r)/EcoRI(SI)
new NarI
amy
SstI
EcoRI
pIJ449
HindIII
ermE
KpnI
SstI
EcoRI
598 b.p.
EcoRI-HindIII
mp10/EcoRI-HindIII
SC65 primer
pSYC1303
RFDNA
new SphI
HindIII
ermE
SstI
EcoRI
small EcoRI-SphI(t)
large
EcoRI-SstI
pUC13
Ap
EcoRI
SstI
HindIII
SphI(t)/EcoRI(SI)
pSYC1311
new NarI
HindIII
amy
ermE
SstI
KpnI
EcoRI

FIG._5.

FIG._6-1.

FIG._6-2.

FIG._6.

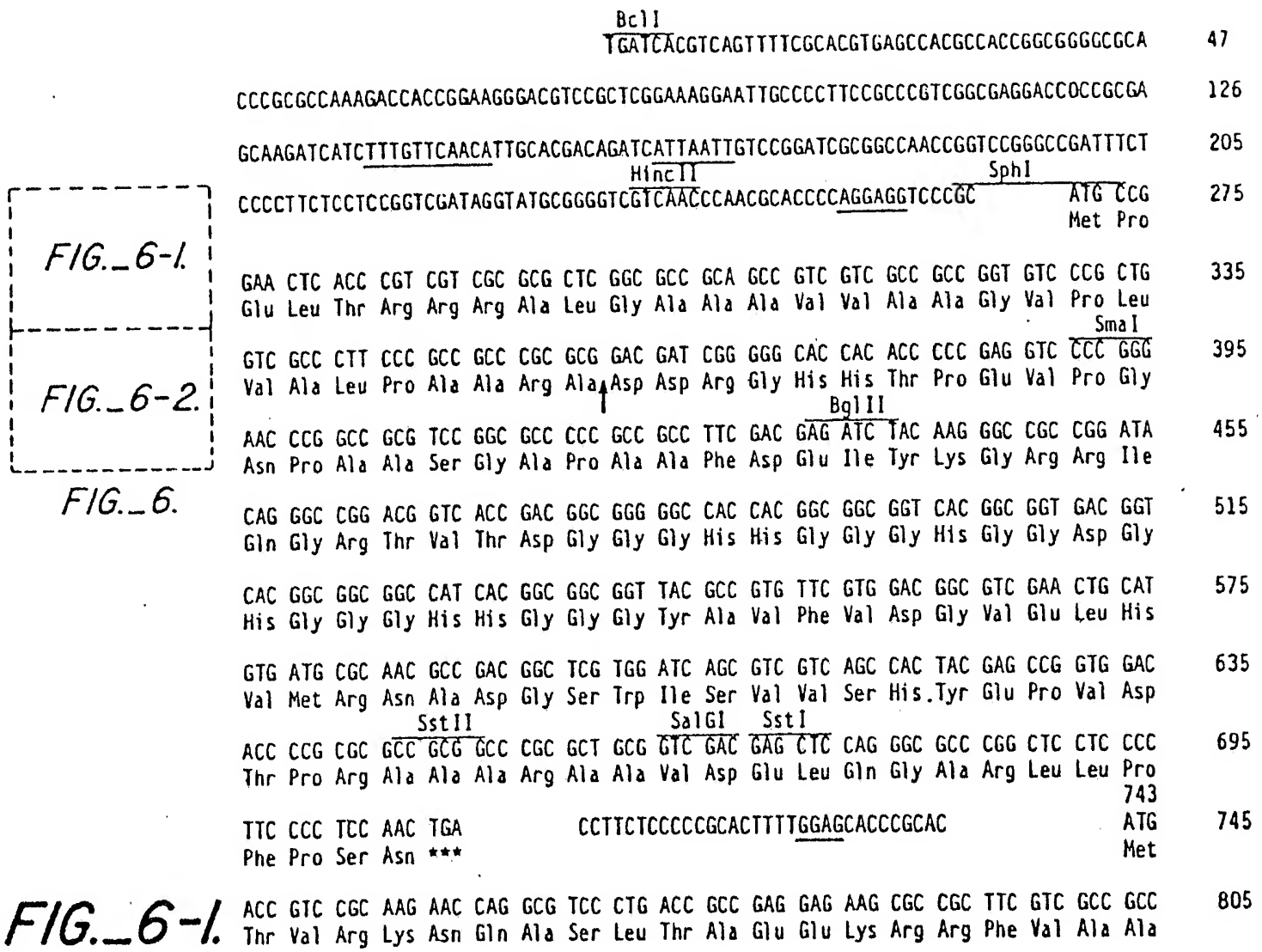


FIG._6-1.

```
CTG CTC GAA CTC AAG CGC ACC GGC CGC TAC GAC GCC TTC GTC ACC ACG CAC AAC GCG TTC      865
Leu Leu Glu Leu Lys Arg Thr Gly Arg Tyr Asp Ala Phe Val Thr Thr His Asn Ala Phe

ATC CTG GGC GAC ACC GAC AAC GGC GAG CGC ACC GGC CAC CGT TCG CCG TCC TTC CTG CCC      925
Ile Leu Gly Asp Thr Asp Asn Gly Glu Arg Thr Gly His Arg Ser Pro Ser Phe Leu Pro

TGG CAC CGC AGA TTT CTG CTG GAG TTC GAG CGG GCG CTC CAG TCG GTG GAC GCG TCG GTG      985
Trp His Arg Arg Phe Leu Leu Glu Phe Glu Arg Ala Leu Gln Ser Val Asp Ala Ser Val

GCG CTG CCG TAC TGG GAC TGG TCC GCC GAC CGG TCC ACC CGG TCC TCG CTG TGG GCG CCG     1045
Ala Leu Pro Tyr Trp Asp Trp Ser Ala Asp Arg Ser Thr Arg Ser Ser Leu Trp Ala Pro

GAC TTC CTC GGC GGC ACC GGG CGC AGC CGG GAC GGC CAG GTG ATG GAC GGG CCG TTC GCC     1105
Asp Phe Leu Gly Gly Thr Gly Arg Ser Arg Asp Gly Gln Val Met Asp Gly Pro Phe Ala

GCG TCG GCC GGC AAC TGG CCG ATC AAT GTG CGG GTG GAC GGC CGT ACG TTC CTG CGG CGG     1165
Ala Ser Ala Gly Asn Trp Pro Ile Asn Val Arg Val Asp Gly Arg Thr Phe Leu Arg Arg

                                                        SalGI
GCG CTC GGC GCG GGC GTG AGC GAA CTG CCC ACG CGT GCC GAG GTC GAC TCG GTG CTG GCG     1225
Ala Leu Gly Ala Gly Val Ser Glu Leu Pro Thr Arg Ala Glu Val Asp Ser Val Leu Ala

ATG GCG ACG TAC GAC ATG GCG CCC TGG AAC AGC GGC TCC GAC GGC TTC CGC AAC CAT CTC     1285
Met Ala Thr Tyr Asp Met Ala Pro Trp Asn Ser Gly Ser Asp Gly Phe Arg Asn His Leu

GAA GGG TGG CGC GGG GTC AAT CTG CAC AAC CGG GTG CAT GTC TGG GTC GGC GGC CAG ATG     1345
Glu Gly Trp Arg Gly Val Asn Leu His Asn Arg Val His Val Trp Val Gly Gly Gln Met

GCG ACC GGG GTC TCC CCC AAC GAC CCG GTG TTC TGG CTG CAC CAC GCC TAC ATC GAC AAG     1405
Ala Thr Gly Val Ser Pro Asn Asp Pro Val Phe Trp Leu His His Ala Tyr Ile Asp Lys

CTG TGG GCC GAG TGG CAG CGG CGG CAC CCC TCG TCC CCG TAT CTG CCG GGC GGC GGC ACG     1465
Leu Trp Ala Glu Trp Gln Arg Arg His Pro Ser Ser Pro Tyr Leu Pro Gly Gly Gly Thr

          SalGI
CCG AAC GTC GTC GAC CTC AAC GAG ACG ATG AAG CCG TGG AAC GAC ACC ACC CCG GCG GCC     1525
Pro Asn Val Val Asp Leu Asn Glu Thr Met Lys Pro Trp Asn Asp Thr Thr Pro Ala Ala

                                                    BclI
CTG CTG GAC CAC ACC CGG CAC TAC ACC TTC GAC GTC TGA TCA                             1567
Leu Leu Asp His Thr Arg His Tyr Thr Phe Asp Val ***
```

FIG._6-2.

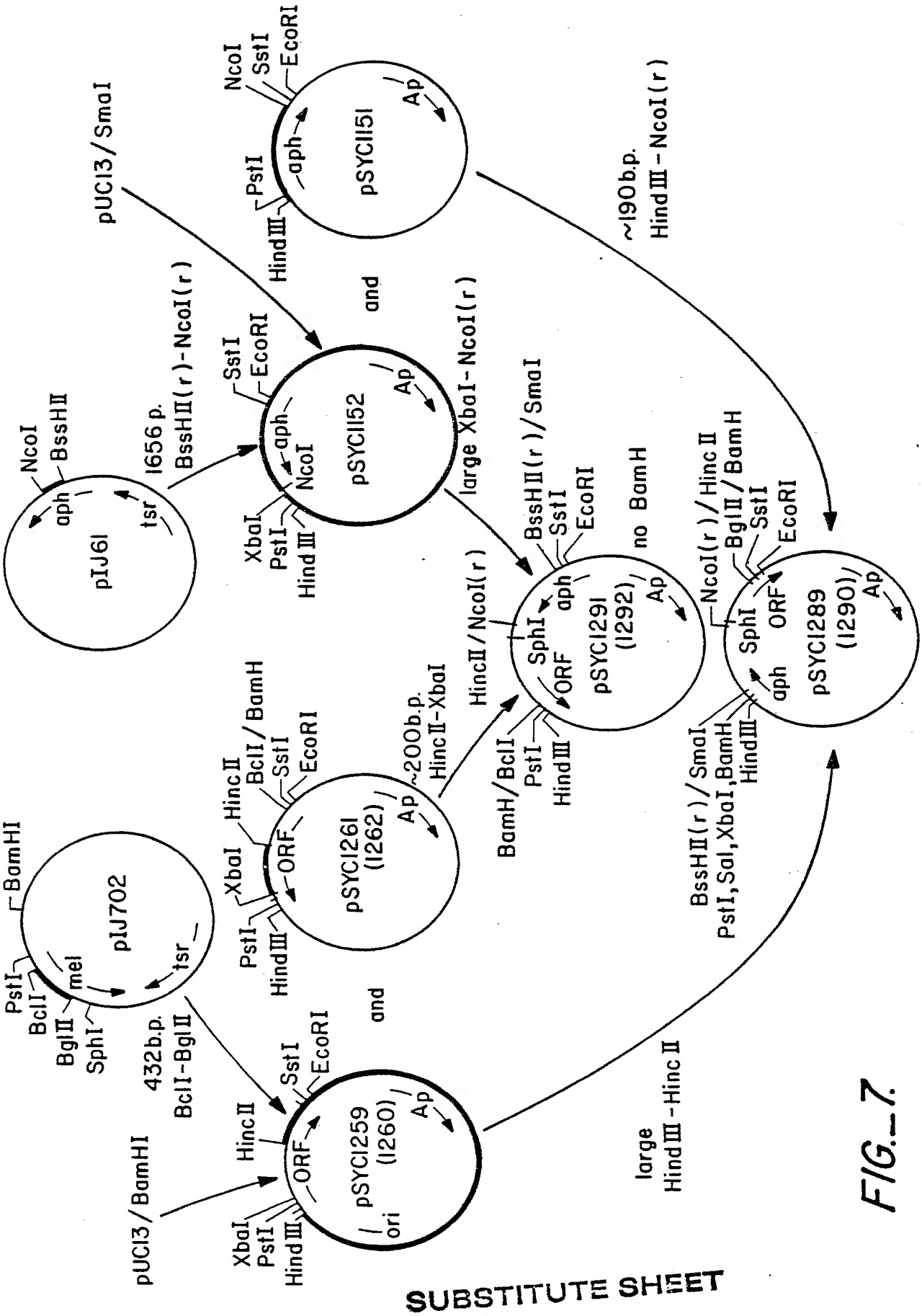

pUC13/SmaI
pUC13/BamHI
pIJ61
pIJ702
pSYC1259 (1260)
pSYC1261 (1262)
pSYC1151
pSYC1152
pSYC1291 (1292)
pSYC1289 (1290)
NcoI
BssHII
aph
tsr
1656 p.
BssHII(r)-NcoI(r)
SstI
EcoRI
Ap
PstI
HindIII
and
large XbaI-NcoI(r)
~190b.p.
HindIII-NcoI(r)
BamHI
PstI
BclI
BglII
SphI
mel
432b.p.
BclI-BglII
HincII
ORF
ori
XbaI
BclI/BamH
~200b.p.
HincII-XbaI
HincII/NcoI(r)
BssHII(r)/SmaI
no BamH
BamH/BclI
NcoI(r)/HincII
BglII/BamH
BssHII(r)/SmaI
PstI,SalI,XbaI,BamH
large
HindIII-HincII

FIG._7.

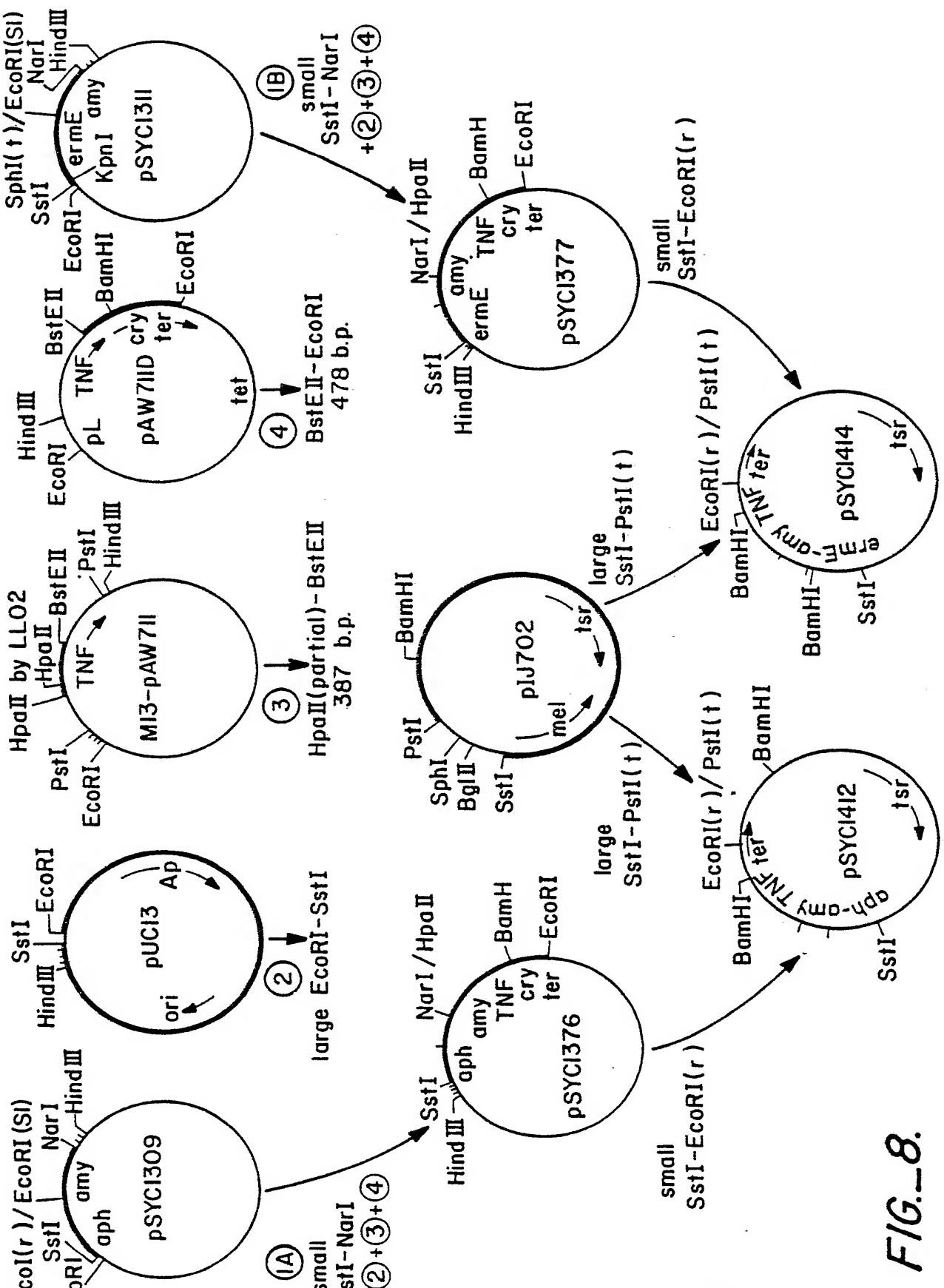


FIG._8.

SUBSTITUTE SHEET

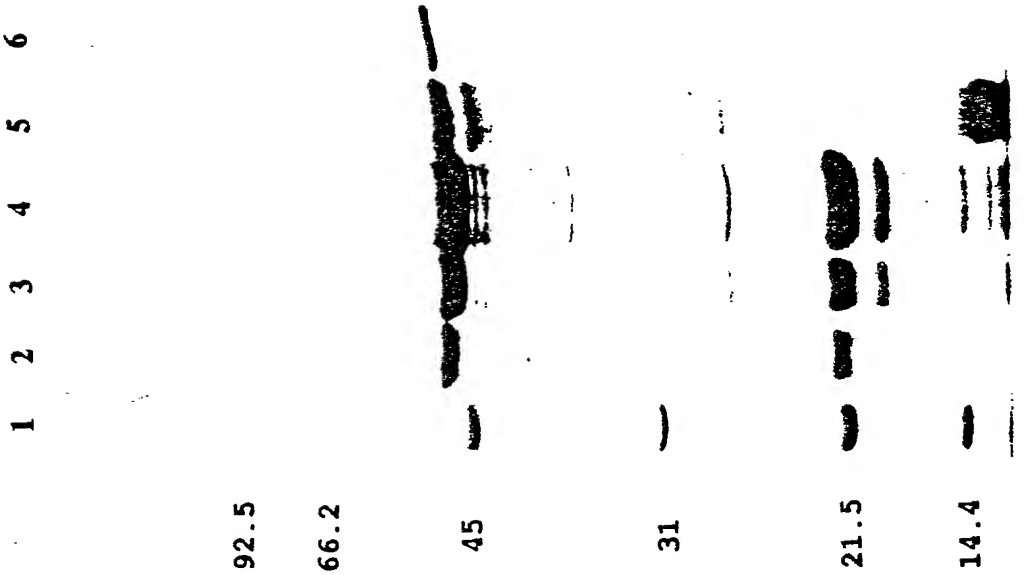


FIG._11.

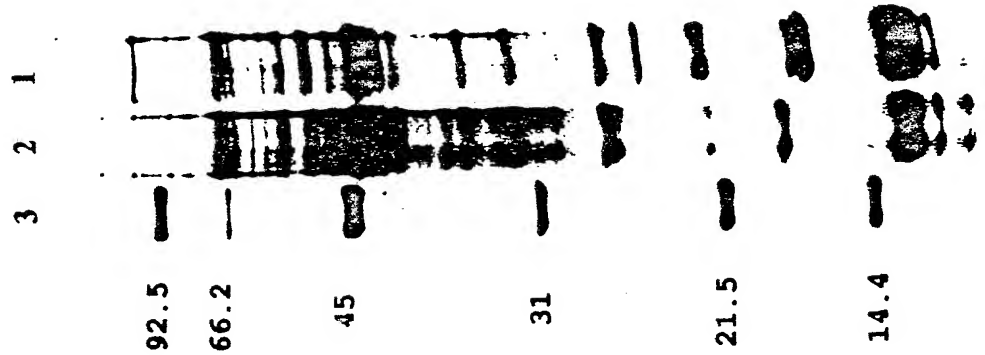


FIG._9.

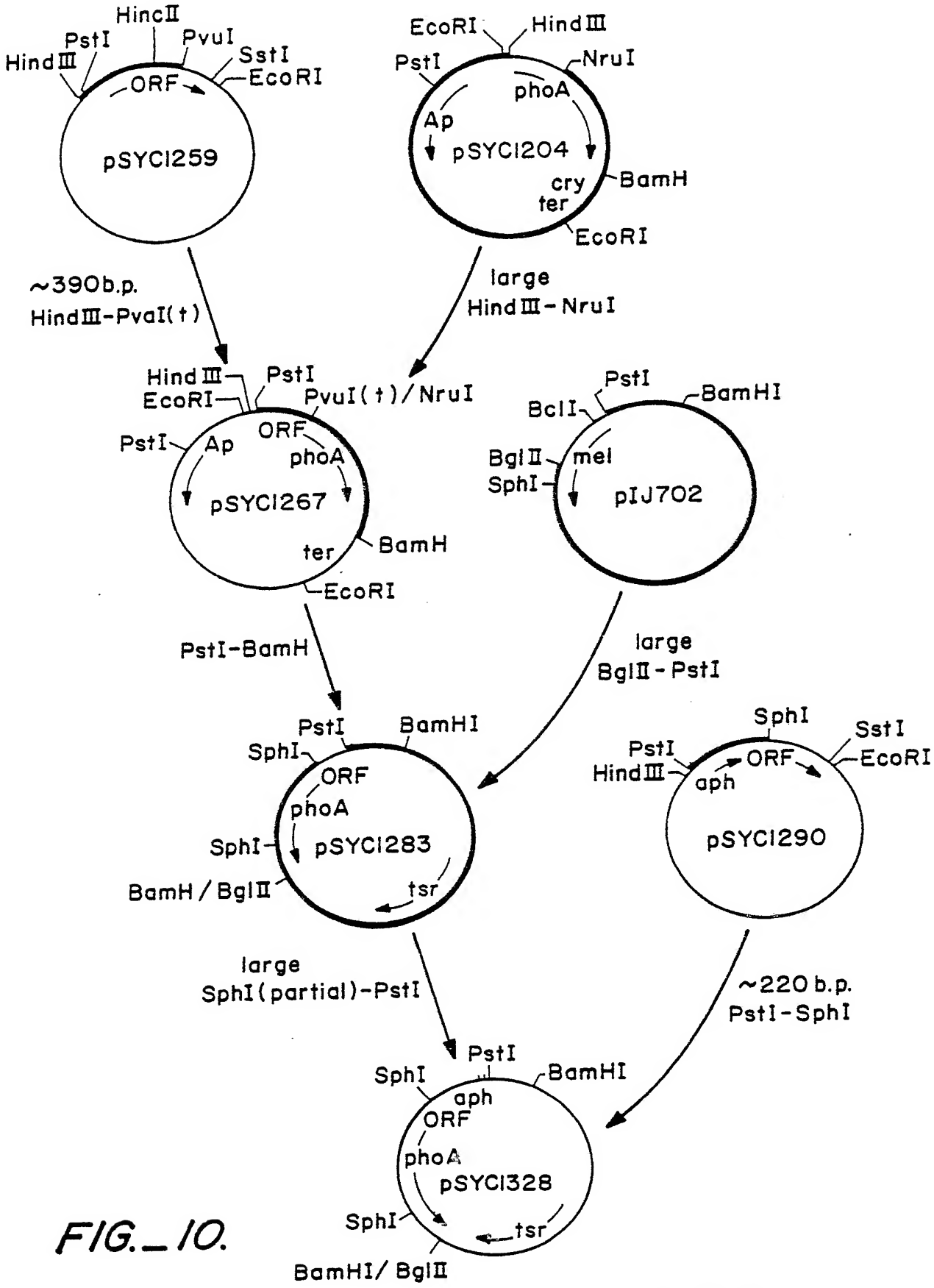


FIG._10.

SUBSTITUTE SHEET

FIG. 12-1. FIG. 12-2.

FIG._12.

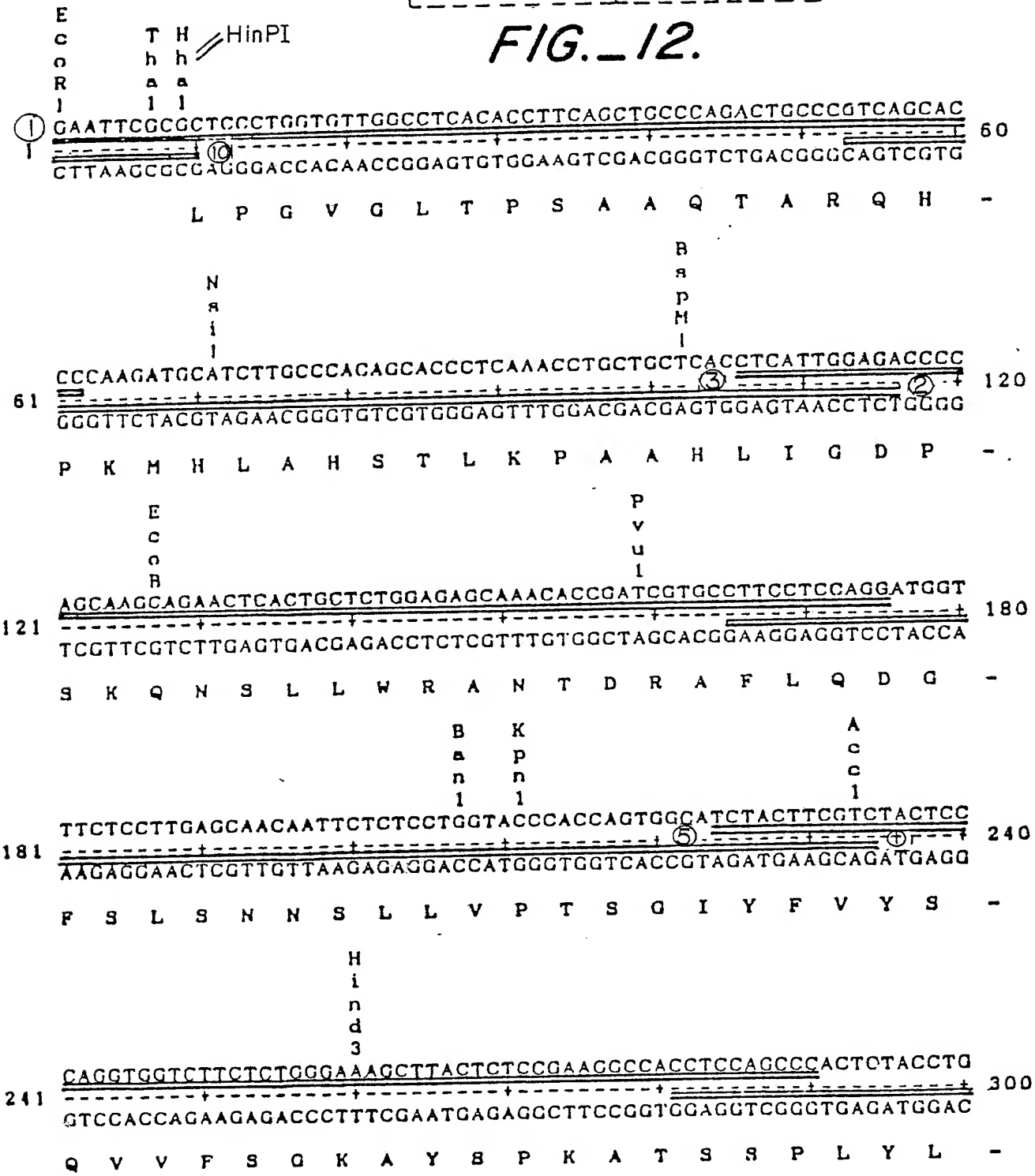


FIG._12-1.

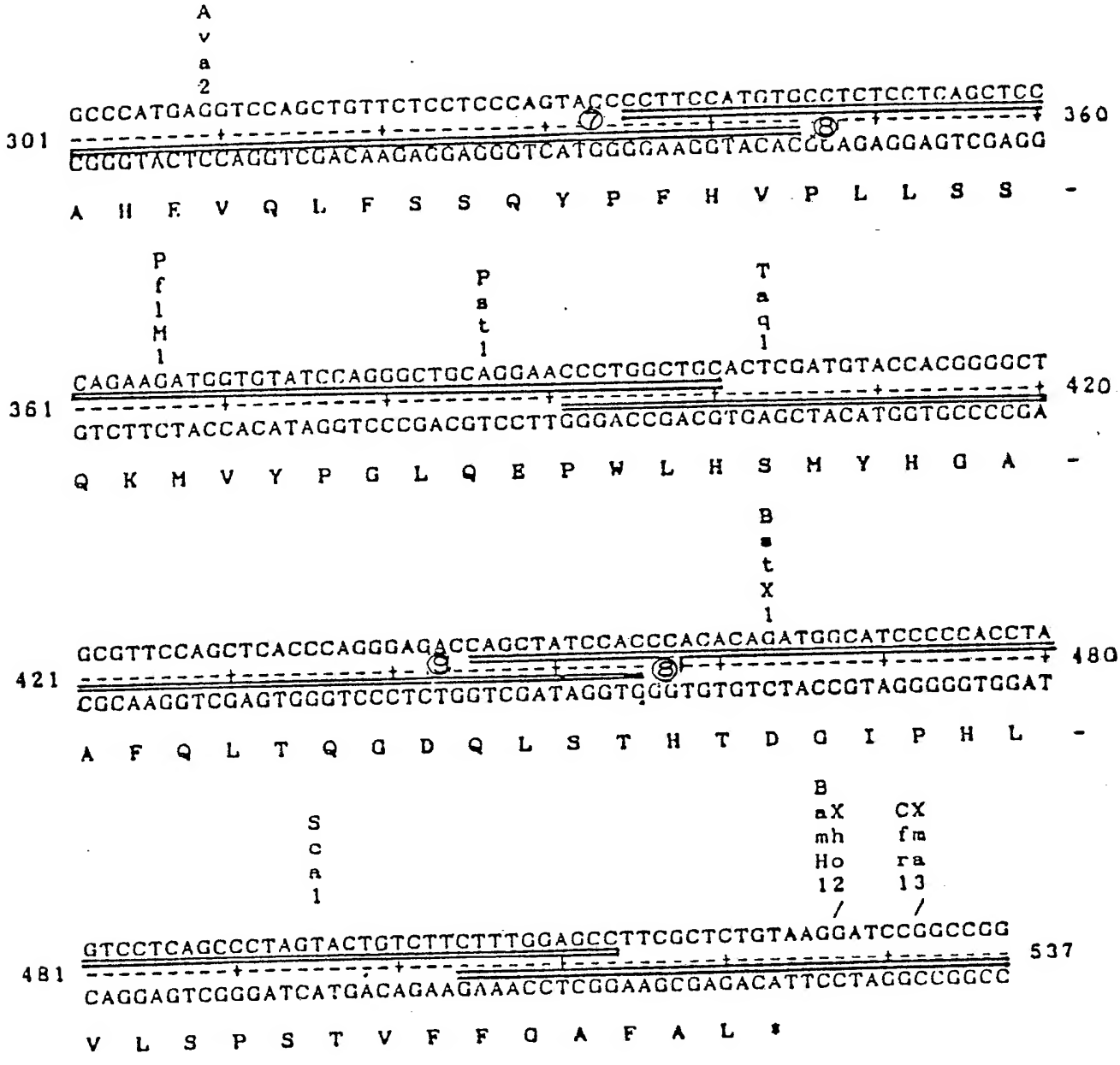


FIG._12-2.

SUBSTITUTE SHEET

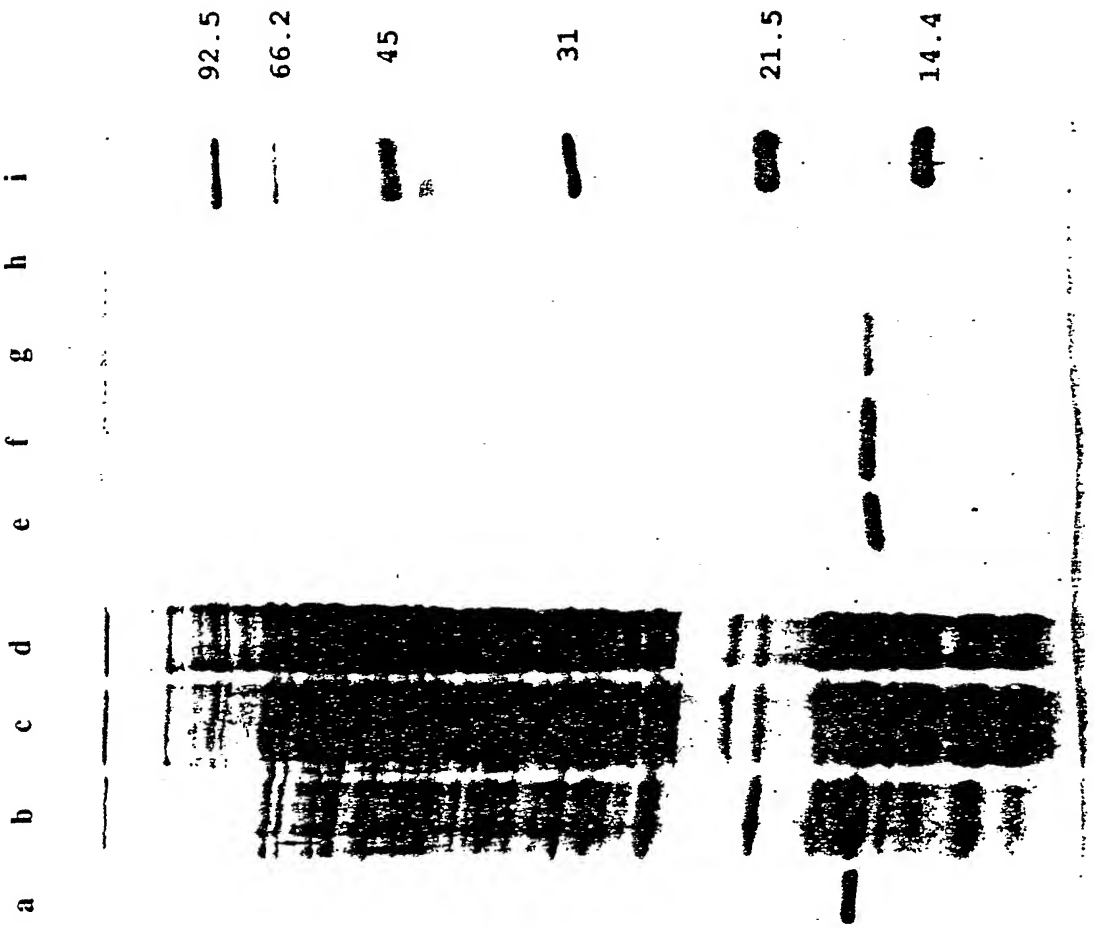


FIG._16.

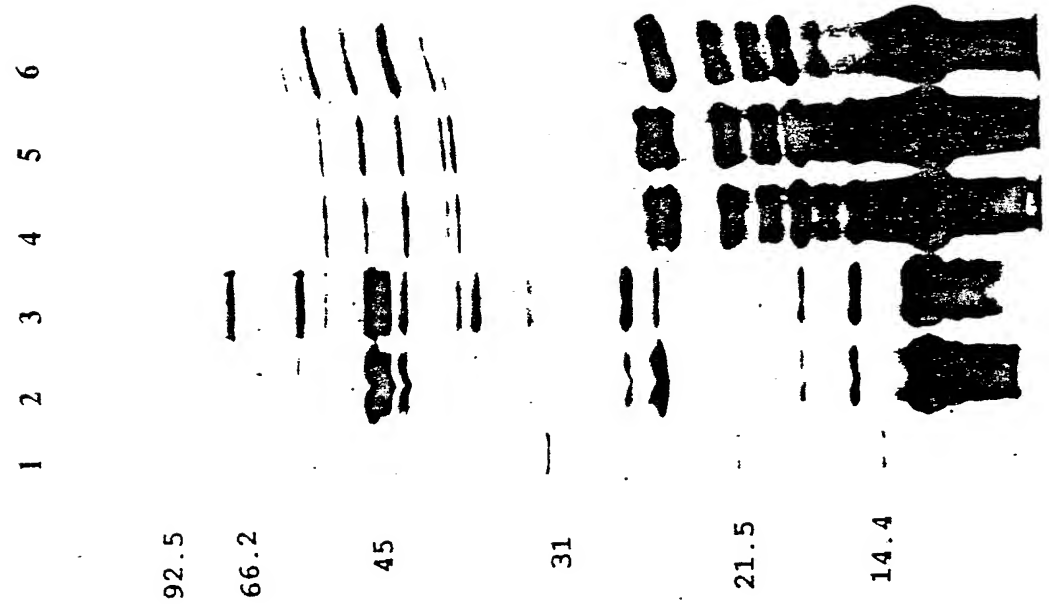


FIG._13.

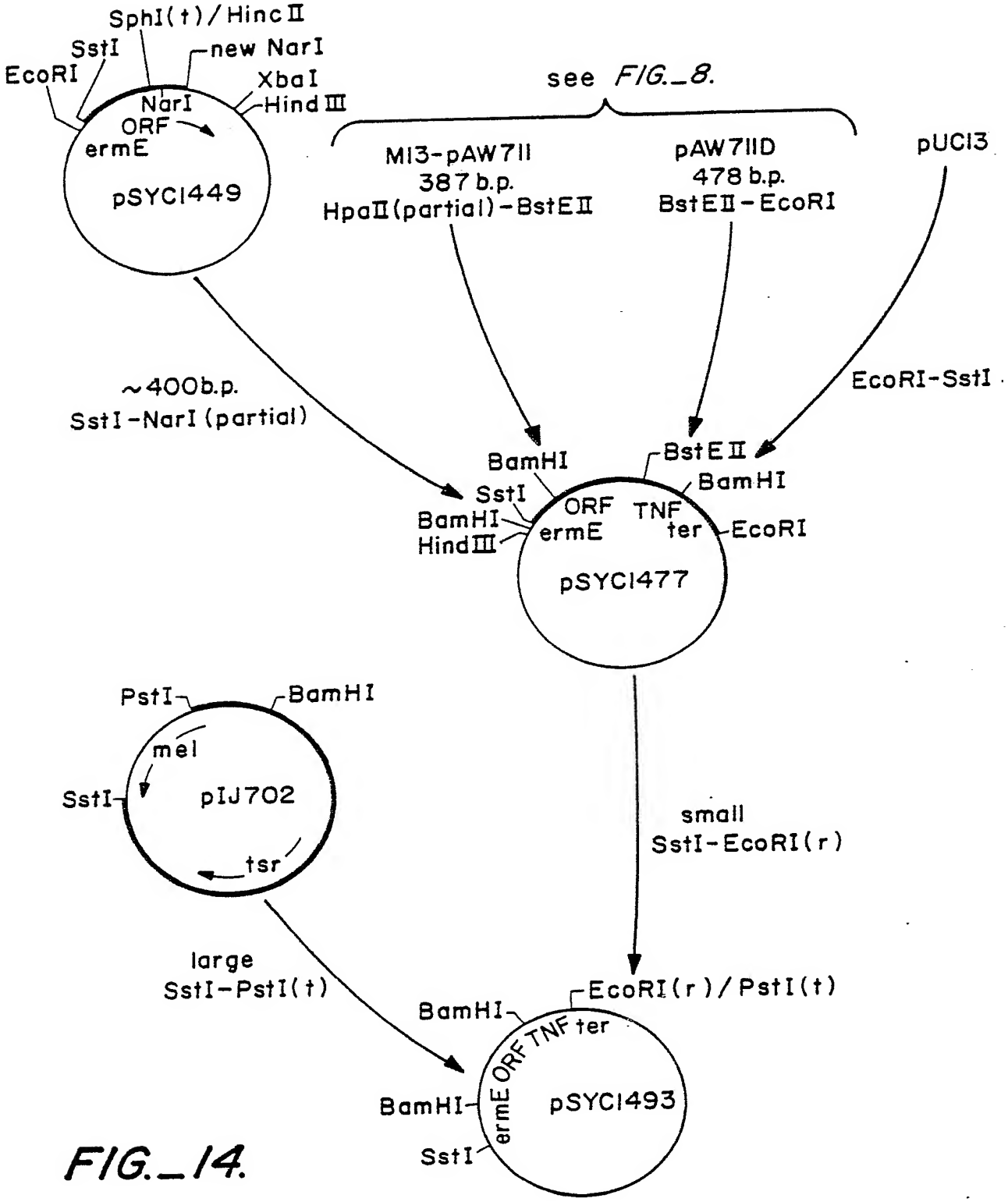

SphI(t)/HincII
SstI
new NarI
EcoRI
XbaI
NarI
HindIII
ORF
ermE
pSYC1449
see FIG._8.
M13-pAW711
387 b.p.
HpaII(partial)-BstEII
pAW711D
478 b.p.
BstEII-EcoRI
pUC13
~400b.p.
SstI-NarI(partial)
EcoRI-SstI
BamHI
BstEII
SstI
BamHI
ORF
TNF
BamHI
ermE
ter
EcoRI
HindIII
pSYC1477
PstI
BamHI
mel
SstI
pIJ702
tsr
small
SstI-EcoRI(r)
large
SstI-PstI(t)
EcoRI(r)/PstI(t)
BamHI
ermE ORF TNF ter
BamHI
pSYC1493
SstI

FIG._14.

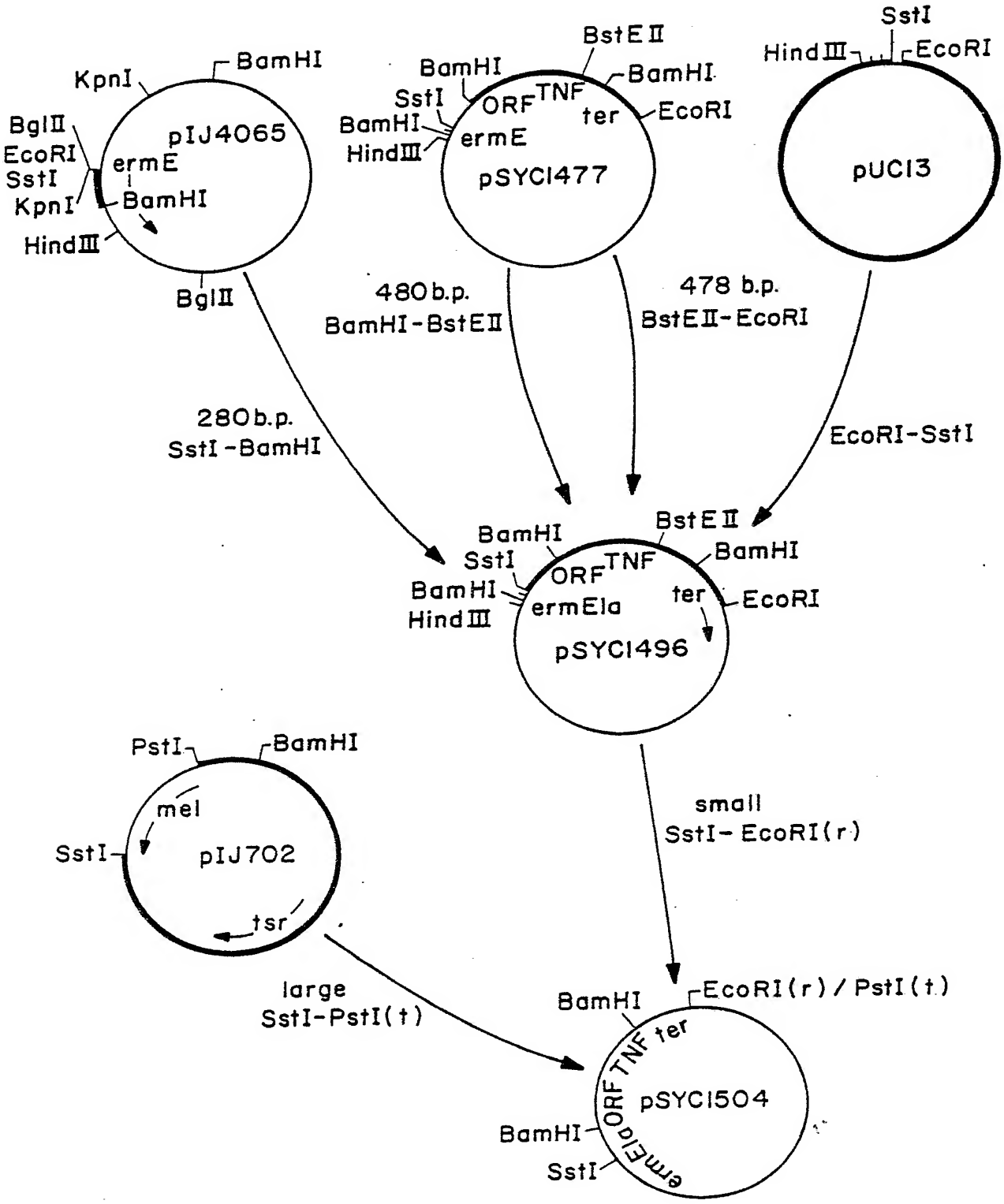

KpnI
BamHI
BglII
EcoRI
SstI
KpnI
ermE
BamHI
HindIII
BglII
pIJ4065
BamHI
BstEII
BamHI
SstI
ORF
TNF
ter
EcoRI
BamHI
HindIII
ermE
pSYC1477
SstI
HindIII
EcoRI
pUC13
480b.p.
BamHI-BstEII
478 b.p.
BstEII-EcoRI
280b.p.
SstI-BamHI
EcoRI-SstI
BamHI
BstEII
SstI
ORF
TNF
BamHI
BamHI
HindIII
ermEla
ter
EcoRI
pSYC1496
PstI
BamHI
mel
SstI
pIJ702
tsr
small
SstI-EcoRI(r)
large
SstI-PstI(t)
BamHI
EcoRI(r)/PstI(t)
ORF
TNF
ter
ermEld
pSYC1504
BamHI
SstI

FIG._15.

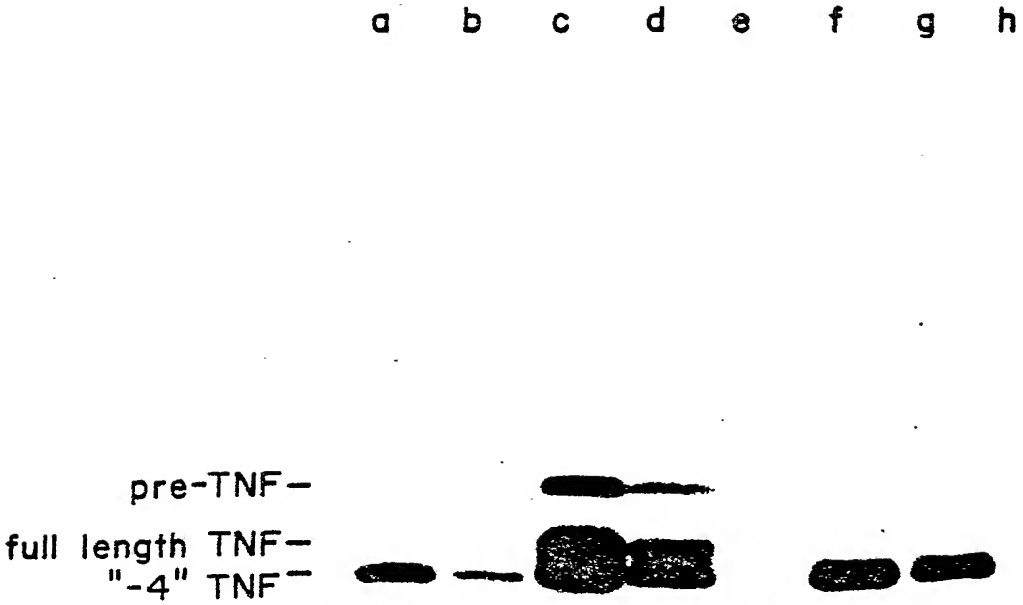


FIG._17.

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 88/00374

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00; C 12 N 1/20

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[4]	C 12 N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

—

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	EP, A, 0148552 (BIOGEN N.V.) 17 July 1985 see page 8, lines 8-25; page 15, line 13 - page 16, line 23; page 9, lines 19-24; page 13, line 30 - page 14, line 10; claim 4; page 9, lines 8-15 cited in the application	1-11
X	EP, A, 0179449 (BIOTECHNICA INTERNATIONAL) 30 April 1986 see the whole document	1-11
X	EP, A, 0196375 (MEIJI SEIKA K.K.) 8 October 1986 see page 5, line 6 - page 8, line 4	1,2,4-6,9-11
P,X	EP, A, 0222279 (CANGENE CORP.) 20 May 1987 see figure 2; page 10, lines 10-16	1,2,4,9-11
P,X	Biological Abstracts/RRM, (Philadelphia, US), S. Chang: "Streptomyces system for producing secreted proteins", ./.	1-11

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
18th April 1988	- 2 JUN 1988
International Searching Authority: EUROPEAN PATENT OFFICE	Signature of Authorized Officer: P.C.G. VAN DER PUTTEN

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	see abstract 33105445, & Abstr. Pap. Am. Chem. Soc., 1987, vol. 194, no. 0, pMBTD 100 --	
A	Gene, volume 32, no. 1/2, December 1984, Elsevier Science Publishers, (Amsterdam, NL), G. Gray et al.: "Synthesis of bovine growth hormone by Streptomyces lividans", see pages 21-30 cited in the application --	3,6,7
A	Journal of General Microbiology, volume 129, part 9, September 1983, SGM, (GB), E. Katz et al.: "Cloning and expression of the tyrosinase gene from Streptomyces antibioticus in Streptomyces lividans", see pages 2703-2714 cited in the application --	
A	Gene, volume 38, no. 1/3, 1985, Elsevier Science Publishers, (Amsterdam, NL), M.J. Bibb et al.: "Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of Streptomyces erythraeus", pages 215-226 see figure 2 cited in the application --	3,8
A	Gene, volume 37, 1985, Elsevier Science Publishers, (Amsterdam, NL), V. Bernan et al.: "The nucleotide sequence of the tyrosinase gene from Streptomyces antibioticus and characterization of the gene product", pages 101-110 see figure 2 cited in the application --	5-8
X	EP, A, 0187630 (SMITHKLINE BECKMAN CORP.) 16 July 1986 see claim 20 --------	1,2,4,9-11

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. US 8800374 SA 20844

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on 17/05/88
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0148552	17-07-85	AU-A- 2991884 JP-A- 60186289	10-01-85 21-09-85
EP-A- 0179449	30-04-86	JP-A- 61158782	18-07-86
EP-A- 0196375	08-10-86	None	
EP-A- 0222279	20-05-87	None	
EP-A- 0187630	16-07-86	US-A- 4717666	05-01-88